

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Ed Vearr - early</u>	NA Sequence (#) <u>2</u>	STN _____
Searcher Phone #: <u>308-45810</u>	AA Sequence (#) _____	Dialog _____
Searcher Location: <u>Biotec lab.</u>	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>9/18/01</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>9/19/01</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>ABSSOI</u>
Clerical Prep Time: <u>4m</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>2m</u>	Other _____	Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:28:43 ; Search time 214.84 Seconds
(without alignments)
5041.567 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtgggtaccacactct.....accacacacacacactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1725	100.0	1725	20 AAX87662	Hepatocyte growth
2	996.6	57.8	1692	20 AAX87663	HGF-MSP hybrid pro
3	857.8	49.7	2190	20 AAX87676	HGF-MSP hybrid pro
4	857.8	49.7	2190	20 AAX87677	HGF-MSP hybrid pro
5	857.8	49.7	2199	12 AAQ15177	Human leukocyte-de
6	857.8	49.7	2199	19 AAV53627	Human leukocyte-de
7	857.8	49.7	2289	11 AAQ06088	Tumour cytotoxic f
8	856.2	49.6	2172	19 AAV61952	Nucleotide sequenc
9	856.2	49.6	2173	14 AAQ37308	Encodes haematopoi
10	856.2	49.6	2289	13 AAQ21066	TCF II-encoding se
11	851.6	49.4	2172	12 AAQ14182	Plasminogen-like g

12	832.8	48.3	2187	14 AAQ47832	Competative inhibi
13	832.8	48.3	2187	14 AAQ47833	Competative inhibi
14	832.8	48.3	2187	14 AAQ45702	Vasculer endotheli
15	832.8	48.3	2214	12 AAQ15176	Human leukocyte-de
16	832.8	48.3	2214	19 AAV53626	Human leukocyte-de
17	831.2	48.2	2184	13 AAQ20049	Human hepatocyte g
18	831.2	48.2	2184	14 AAQ46040	Hepatocyte growth
19	831.2	48.2	2187	12 AAQ10489	Hepatic parenchyma
20	829.6	48.1	2393	12 AAQ14038	Human hepatocyte g
21	828	48.0	2184	15 AAQ56157	Hepatocyte growth
22	825	47.8	873	13 AAQ23662	Truncated hepatocy
23	825	47.8	873	14 AAQ34613	cDNA encoding 34kd
24	825	47.8	873	18 AAT78409	34 kilodalton hepa
25	821.6	47.6	2184	13 AAQ26052	Human HGF. Homo s
26	820	47.5	1661	12 AAQ12398	Human hepatocyte g
27	820	47.5	1661	13 AAQ22144	Human HGF gene par
28	820	47.5	2187	12 AAQ12399	Human hepatocyte g
29	820	47.5	2187	13 AAQ22146	Complete human HGF
30	820	47.5	2187	19 AAV20546	Human recombinant
31	820	47.5	2187	20 AAX15186	Nucleic acid encod
32	682.4	39.6	1400	13 AAQ20745	Encodes alpha-chai
33	680.8	39.5	2187	13 AAQ20255	Encodes alpha- and
34	585	33.9	648	18 AAT78410	Hepatocyte growth
35	302.4	17.5	2394	13 AAQ22143	Human HGF gene par
36	302.4	17.5	2395	12 AAQ12397	Human hepatocyte g
37	170.6	9.9	366	13 AAQ26714	Encodes exon II of
38	159	9.2	2021	15 AAQ79725	Human L5/3 tumour
39	159	9.2	2021	18 AAT62438	Human L5/3 partial
40	159	9.2	2213	18 AAT47149	Macrophage stimula
41	159	9.2	2216	18 AAT47145	Macrophage stimula
42	159	9.2	2216	20 AAV72083	Human MSP cDNA. H
43	159	9.2	2219	15 AAQ79723	Human L5/3 tumour
44	159	9.2	2219	18 AAT62436	Human L5/3 partial
45	159	9.2	2262	15 AAQ79729	Full-length human

ALIGNMENTS

RESULT 1
AXX87662
ID AAX87662 standard; cDNA; 1725 BP.
XX AC AAX87662;
XX AC AAX87662;
DT 26-OCT-1999 (first entry)
XX DE Hepatocyte growth factor hybrid protein Magic F-1 cDNA.
XX KW Magic Factor-1; human; hepatocyte growth factor; HGF; apoptosis;
KW chemotherapy; toxicity; therapy; ss.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT sig_peptide 1..63
FT mat_peptide 64..1725
FT /*tag= a
FT /*tag= b
XX PN WO9938967-A2.
XX PD 05-AUG-1999.
XX PF 27-JAN-1999; 99WO-EP00478.
XX PR 30-JAN-1998; 98IT-MI00179.
XX PA (DOMP-) DOMPE SPA.
XX PI Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;
XX

Db 1561 cccgacaaagggttgatgataattatgcgcgaatcccgatggccagccgagccacg 1620
 Qy 1621 tgcatactcttgacctcacaccgcgtggagtgactgtgcaattaaacatggcgtag 1680
 Db 1621 tgcatactcttgacctcacaccgcgtggagtgactgtgcaattaaacatggcgtag 1680
 Qy 1681 aaagctgacgacgacgacacacacacacacacacacacacacacacacacac 1725
 Db 1681 aaagctgacgacgacgacacacacacacacacacacacacacacacacacac 1725

RESULT 2

AXX87663
 ID AAX87663 standard; cDNA; 1692 BP.

XX AC AAX87663;

XX DT 26-OCN-1999 (first entry)

XX HGF-MSP hybrid protein Metron Factor-1 cDNA.

XX Metron Factor-1; human; hepatocyte growth factor; HGF;
 KW macrophage stimulating protein MSP; apoptosis; chemotherapy;
 KW toxicity; therapy; ss.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT sig_peptide 1..54

FT mat_peptide /*tag= a

FT /*tag= b

XX W09938967-A2.

XX 05-AUG-1999.

XX 27-JAN-1999; 99WO-EP00478.

XX 30-JAN-1998; 98IT-MI00179.

XX (DOMP-) DOMPE SPA.

XX Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;

XX WPI; 1999-494090/41.

XX P-PSDB; AAY06620.

XX Recombinant proteins from recombination of HGF and MSP structural
 PT domains, useful for protection against apoptosis induced by
 PT chemotherapeutics

XX Example 1; Page 60-61; 63pp; English.

XX This DNA sequence codes for Metron F-1 factor (see AAY06620), a
 CC recombinant protein composed of the macrophage stimulating protein
 CC (MSP) alpha chain signal peptide, hairpin loop and kringle
 CC domains 1 and 2 joined via a peptide linker to the hepatocyte
 CC growth factor (HGF) hairpin loop and kringle domains 1 and 2 plus a
 CC polyhistidine tag. The portions of DNA encoding the various
 CC regions of Metron F-1 were obtained by PCR amplification and then
 CC recombined to obtain the hybrid sequence. Expression vectors,
 CC prokaryotic or eukaryotic host cells and a process for preparing
 CC recombinant proteins from HGF and MSP are claimed. Metron F-1 and
 CC other recombinant proteins comprising HGF and/or MSP structural
 CC domains are used to prevent or treat chemotherapeutic-induced
 CC toxicity such as myelotoxicity, hepatotoxicity, nephrotoxicity,
 CC mucotoxicity and neurotoxicity (claimed). They protect cells from
 CC death (apoptosis) induced by chemotherapy of, e.g. tumors. In
 CC particular, they can be used for expansion of marrow precursors, to
 CC increase proliferation of the haematopoietic precursors or to

CC stimulate their entry in the cycle. By modification of the
 CC proteolytic site, hybrid factors can be obtained which are activated
 CC by proteases of the endoplasmic reticulum (such as furines) during
 CC their synthesis. When the proteolytic site is removed, permanently
 CC immature forms of the factors can be obtained, having a potential
 CC partial agonistic or antagonistic activity. Different functional
 CC domains can be combined so as to modulate the biological effects.

XX Sequence 1692 BP; 442 A; 457 C; 459 G; 334 T; 0 other;

Query Match 57.8%; Score 996.6; DB 20; Length 1692;
 Best Local Similarity 83.2%; Pred. No. 4e-249;
 Matches 1166; Conservative 0; Mismatches 214; Indels 21; Gaps 2;

Qy 346 ttgacctctatgaaacaaagactacattagaaactgcacatctggttaagagcagcag 405
 Db 292 tgtgacctctccagaagaagactacgtacggacctgcacatgaacaaatgggttggg 351
 Qy 406 tacaagggaacagtgatctatcataagagtggtcatcaaatgtcagccctggagtccatg 465
 Db 352 taccggggcaccatggccagaccgtgggtgcctgcctgcctgcctggagcacaag 411
 Qy 466 ataccacacgaacacagcta-----tcgggggtaaagacctacaggaataactctgt 516
 Db 412 ttcccgaaatgacacaagtacacgcccactctccggaatggcctggaagaacttctgc 471
 Qy 517 cgaatactctcgaggggaagaaggacacctgtgtttcacagaacacacagaggtacgc 576
 Db 472 cgtaaccttgatggcgaaccccgaggctctgtgtgtacacacacacacacctgtgtgcgc 531
 Qy 577 tacgaagtctgtgacattctcctcagtggttcagaagttggaatgcacacacacacacac 636
 Db 532 ttccagagtcgagcatcaaatcctgcgggagcgccgtgtgtgtgtgtgcaatggcgag 591
 Qy 637 agttatcaggtctcatggatcacacagaatcagcgaagattgtcagcgcgtggatcat 696
 Db 592 gaataccgcggcggtagaccgcagcggagtcagggcgagtcgacgcgtggatctt 651
 Qy 697 cagacacacacacgcacaaattcttgcctgaaagatatcccgacacagggcttctgatg 756
 Db 652 cagcaccgcgac 711
 Qy 757 aattatgcgcgaatcccgatggccagcgcgagccatgtgtgtatctcttgacctcac 816
 Db 712 aactatgcgcgaatcccgatggccagcgcgagccatgtgtgtatctcttgacctcac 771
 Qy 817 accgcgtggagtgactgtgcaattaaacacacacacacacacacacacacacacacac 864
 Db 772 atcgagcagagagttctgtgacctccccctgcgggtccgagcagacacacacacacac 831
 Qy 865 ggcgtggcgttctgtgtggcgtgtgcctccgcggtggcgttctctagaggacacaaag 924
 Db 832 ggcgtggcgttctgtgtggcgtgtgcctccgcggtggcgttctctagaggacacaaag 891
 Qy 925 aaaaagaataacataatcatgaattcaaaaaatcagaaagactaccctaatcaaaaa 984
 Db 892 aaaaagaataacataatcatgaattcaaaaaatcagaaagactaccctaatcaaaaa 951
 Qy 985 gatccagcactgaagataaaaaacccaaagtgaaatcactgcagacacacacacacacac 1044
 Db 952 gatccagcactgaagataaaaaacccaaagtgaaatcactgcagacacacacacacacac 1011
 Qy 1045 tttactagaataaaaggacttcattcactcgaagctttgttttttataaagaaga 1104
 Db 1012 tttactagaataaaaggacttcattcactcgaagctttgttttttataaagaaga 1071
 Qy 1105 aaacaatgcctctgttcccttccatagcatgtcgaagtgagtgagtgagtgagtgagtgag 1164
 Db 1072 aaacaatgcctctgttcccttccatagcatgtcgaagtgagtgagtgagtgagtgag 1131
 Qy 1165 catgaattgacctctatgaaacaaagactacattagaactgcattcattgtglaaagga 1224

Qy 421 tctatcactaagagtggaatcaaatgtcagccctggagttccatgataaccacgaacac 480
 |||||||
 Db 421 tctatcactaagagtggaatcaaatgtcagccctggagttccatgataaccacgaacac 480
 |||||||
 Qy 481 agctatcggttaagacctacagaaaactactgtcgaaatcctcgagggaagagg 540
 |||||||
 Db 481 agctatcggttaagacctacagaaaactactgtcgaaatcctcgagggaagagg 540
 |||||||
 Qy 541 ggaacctgtgtttcacagcaatccagaggtacgctacgaagtctgtgacattctcag 600
 |||||||
 Db 541 ggaacctgtgtttcacagcaatccagaggtacgctacgaagtctgtgacattctcag 600
 |||||||
 Qy 601 ttttcaagaagttgaatgcacgtgcaatgggagagttatcgaggtctcatgataat 660
 |||||||
 Db 601 ttttcaagaagttgaatgcacgtgcaatgggagagttatcgaggtctcatgataat 660
 |||||||
 Qy 661 acagaatacaggaagattgtcagcgtggtatcatcagacacacacacgcgcaaaattc 720
 |||||||
 Db 661 acagaatacaggaagattgtcagcgtggtatcatcagacacacacacgcgcaaaattc 720
 |||||||
 Qy 721 ttgctgaagaatcccgacagggctttgatgataattatgccgaatcccgatggc 780
 |||||||
 Db 721 ttgctgaagaatcccgacagggctttgatgataattatgccgaatcccgatggc 780
 |||||||
 Qy 781 cagccgagggccatggtctatactcttgaccctcacaccccgctggagtgactgtgaatt 840
 |||||||
 Db 781 cagccgagggccatggtctatactcttgaccctcacaccccgctggagtgactgtgaatt 840
 |||||||
 Qy 841 aaaacatgcgtgacaaagct 861
 |||||||
 Db 841 aaaacatgcgtgacaaatct 861
 |||||||

RESULT 4

AAX87677
 ID AAX87677 standard; cDNA; 2190 BP.

AC AAX87677;
 XX AC
 XX AC
 DT 26-OCT-1999 (first entry)
 XX DT
 XX DT
 DE HGF-MSP hybrid protein alphabeta-RTKR factor cDNA.
 XX DE
 XX DE
 KW Alphabeta-RTKR; human; hepatocyte growth factor; HGF; MSP;
 KW macrophage stimulating protein; apoptosis; chemotherapy; toxicity;
 KW therapy; ss.
 XX KW
 OS Homo sapiens.
 OS Synthetic.
 XX OS
 XX OS
 FT key Location/Qualifiers
 FT sig_peptide 1..63
 FT /*tag= a
 FT mat_peptide 63..2190
 FT /*tag= b
 XX FT
 XX FT
 PN W09938968-A1.
 XX PN
 XX PN
 PD 05-AUG-1999.
 XX PD
 XX PD
 PF 28-JAN-1999; 99WO-EP00502.
 XX PF
 XX PF
 PR 30-JAN-1998; 98IT-MI00180.
 XX PR
 XX PR
 PA -(DMP-) DOMPE SPA.
 XX PA
 XX PI
 PI Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;
 XX PI
 XX PI
 DR WPI: 1999-494091/41.
 DR -p-PSDB; AAY06622.
 XX DR
 PT Recombinant proteins from recombination of HGF and MSP structural

domains, useful for protecting cells against apoptosis induced by chemotherapeutics

Example 2; Page 56-58; 63pp; English.

CC This DNA sequence codes for alphabeta-RTKR factor (see AAY06622), a
 CC recombinant protein composed of the hepatocyte growth factor (HGF)
 CC alpha chain (i.e. the signal sequence, hairpin loop and kringles
 CC 1-4), the natural cleavage site of HGF plus a furin protease
 CC cleavage site, the macrophage stimulating protein (MSP) beta chain
 CC and a polyhistidine tag sequence. This structure allows the
 CC recombinant protein to interact with both the HGF receptor (Met)
 CC and the MSP receptor (Ron) and thereby induce biological responses
 CC which are synergistic and selective compared with the natural
 CC factors and truncated forms of the proteins. The portions of DNA
 CC encoding the various portions of alphabeta-RTKR are obtained by PCR
 CC amplification of HGF or MSP cDNA and then recombined to obtain the
 CC hybrid sequence. Expression vectors, prokaryotic or eukaryotic
 CC host cells and a process for preparing recombinant proteins from
 CC HGF and MSP are claimed. Alphabeta-RTKR and other HGF-MSP
 CC recombinant proteins are used for the prevention or treatment of
 CC chemotherapeutic-induced toxicity such as myelotoxicity,
 CC hepatotoxicity, nephrotoxicity, mucotoxicity and neurotoxicity
 CC (claimed). They protect cells from programmed cell death
 CC (apoptosis) induced by chemotherapy of, e.g. tumors. In particular,
 CC they can be used for expansion of marrow precursors, to increase
 CC proliferation of the haematopoietic precursors or to stimulate
 CC their entry in the cycle. By modification of the proteolytic
 CC site, hybrid factors can be obtained which are activated by
 CC proteases of the endoplasmic reticulum (such as furines) during
 CC their synthesis. When the proteolytic site is removed, permanently
 CC immature forms of the factors can be obtained, having a potential
 CC partial agonistic or antagonistic activity. Different functional
 CC domains can be combined so as to modulate the biological effects.

XX Sequence 2190 BP; 612 A; 528 C; 544 G; 506 T; 0 other;

Query Match 49.7%; Score 857.8; DB 20; Length 2190;

Best Local Similarity 99.8%; Pred. No. 4.8e-213;

Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgtgggtgacaaactcctgcagccctgtctgcagcatgtcctcctgcattcctc 60
 |||||||
 Db 1 atgtgggtgacaaactcctgcagccctgtctgcagcatgtcctcctgcattcctc 60
 |||||||
 Qy 61 ctgctcccatcgccatccctctatgcagagggaagaaagaaagaaatacaattcat 120
 |||||||
 Db 61 ctgctcccatcgccatccctctatgcagagggaagaaagaaagaaatacaattcat 120
 |||||||
 Qy 121 gaattcaaaaaatcagaaagactaccctaaatcaaaatagatccagcactgaagataaaa 180
 |||||||
 Db 121 gaattcaaaaaatcagaaagactaccctaaatcaaaatagatccagcactgaagataaaa 180
 |||||||
 Qy 181 accaaaaaagtgaatactactgcagaccatgtgctaatagtactaggataaaggactt 240
 |||||||
 Db 181 accaaaaaagtgaatactactgcagaccatgtgctaatagtactaggataaaggactt 240
 |||||||
 Qy 241 ccattcaactgcaaggctttgttttgataaagcaagaaacaatgcctctggtccccc 300
 |||||||
 Db 241 ccattcaactgcaaggctttgttttgataaagcaagaaacaatgcctctggtccccc 300
 |||||||
 Qy 301 ttcaatagcatgtcaagtggagtgaataaagaatttggccatgaatttgaccttatgaa 360
 |||||||
 Db 301 ttcaatagcatgtcaagtggagtgaataaagaatttggccatgaatttgaccttatgaa 360
 |||||||
 Qy 361 aacaaagactacattagaaactgcattcattggttaaaggacgcagctacaagggaacagta 420
 |||||||
 Db 361 aacaaagactacattagaaactgcattcattggttaaaggacgcagctacaagggaacagta 420
 |||||||
 Qy 421 tctatcaactaagatggcgtacaaatgtcagccctggagttccatgataccacacgaacac 480
 |||||||
 Db 421 tctatcaactaagatggcgtacaaatgtcagccctggagttccatgataccacacgaacac 480
 |||||||

Db 808 cagccgagccatggtgctatactctgacccctcacaccgcgtggagtgactgtgcaatt 867

Qy 841 aaacatgcgtgcaaaagct 861
 |||||
 Db 868 aaacatgcgtgcaaaact 888

RESULT 6

AAV53627
 ID AAV53627 standard; cDNA; 2199 BP.

XX AC AAV53627;

XX DT 21-DEC-1998 (first entry)

XX Human leukocyte-derived hepatocyte growth factor cDNA clone HLC2.

XX Hepatocyte-derived growth factor; HGF; human; liver; hepatoma;

KW diagnosis; therapy; ss.

XX Homo sapiens.

OS XX

Key Location/Qualifiers
 CDS 28..2199
 FT /*tag= a

XX EP859009-A2.

XX 19-AUG-1998.

XX 07-JUN-1991; 91EP-0109369.

XX 11-JUN-1990; 90JP-0152474.

XX (NAKA/) NAKAMURA T.

XX Asami O, Hagiya M, Ihara I, Nakamura T, Sakaguchi M;

PI Seki T, Shimizu S, Shimonishi M;

XX WPI; 1998-429650/37.

DR P-PSDB; AAW59923.

XX Preparation of recombinant hepatocyte growth factor polypeptide - by
 PT culturing mammalian cells transformed with vector containing human
 PT leukocyte-derived HGF gene

XX Example 1; Fig 3a-c; 30pp; English.

XX cDNA clone HCL2 codes for a human leukocyte-derived hepatocyte
 CC growth factor (HGF, see AAW59923). It was isolated from a human
 CC leukocyte cDNA library using a fragment of human liver-derived HGF
 CC cDNA as probe. Another clone, HLC3 (see AAV53626), was also
 CC isolated. HCL2 differs from HCL3 by having nucleotides 510-524 of
 CC HCL3 deleted. A claimed method of producing a HGF comprises:
 CC transforming mammalian cells (preferably CHO cells) with a
 CC recombinant expression vector that has a promoter (preferably the
 CC SV40 early promoter), a gene coding for human leukocyte-derived
 CC HGF, and the dihydrofolate reductase gene; (b) culturing the
 CC transformed cells in the presence of successively elevated
 CC concentrations of methotrexate; and (c) recovering the HGF
 CC polypeptide from the culture supernatant. HGF polypeptides may be
 CC used as hepatocyte cultivation reagents, liver regeneration
 CC promoters, in basic research on liver function, research on the
 CC action of various hormones and drugs on hepatocytes, research on
 CC the carcinogenesis mechanism of hepatoma, clinical diagnostic
 CC reagents using an antibody against the polypeptide and therapeutic
 CC drugs for liver disease.

XX Sequence 2199 BP; 678 A; 466 C; 503 G; 552 T; 0 other;

Query Match 49.7%; Score 857.8; DB 19; Length 2199;
 Best Local Similarity 99.8%; Pred. No. 4.8e-213;

Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgtgggtgaccaaactctcgcagccctgtgtgctgagcatgtctctgtctctctc 60
 |||||
 Db 28 atgtgggtgaccaaactctcgcagccctgtgtgctgagcatgtctctgtctctc 87
 |||||
 Qy 61 ctgctcccatcgccatccctatcgaggggacaaagaaaagaatacaattcat 120
 |||||
 Db 88 ctgctcccatcgccatccctatcgaggggacaaagaaaagaatacaattcat 147
 |||||
 Qy 121 gaattcaaaaaatcagcaaaagactaccctcaatcaaaatagatccagcactgaagataaaa 180
 |||||
 Db 148 gaattcaaaaaatcagcaaaagactaccctcaatcaaaatagatccagcactgaagataaaa 207
 |||||
 Qy 181 accaaaaagtgaatactcagaccacaaatgtctaataatgtactagggaataaagactt 240
 |||||
 Db 208 accaaaaagtgaatactcagaccacaaatgtctaataatgtactagggaataaagactt 267
 |||||
 Qy 241 ccattcaacttgcaaggctttttttgtataaagcaagaacaaatgcctctggtccccc 300
 |||||
 Db 268 ccattcaacttgcaaggctttttttgtataaagcaagaacaaatgcctctggtccccc 327
 |||||
 Qy 301 ttcaatagcatgtcaagtggagtgaataaagaaatttggccatgaatttgacctctatgaa 360
 |||||
 Db 328 ttcaatagcatgtcaagtggagtgaataaagaaatttggccatgaatttgacctctatgaa 387
 |||||
 Qy 361 acaaaagactacattagaaactgcatttggtaaggagcgcagctacaaagggaacagta 420
 |||||
 Db 388 acaaaagactacattagaaactgcatttggtaaggagcgcagctacaaagggaacagta 447
 |||||
 Qy 421 tctatcaactaagagtggcatcaaaatgcagccctgaggttccatgataccacgaacac 480
 |||||
 Db 448 tctatcaactaagagtggcatcaaaatgcagccctgaggttccatgataccacgaacac 507
 |||||
 Qy 481 agctatcgggttaaagacacctacaggaactactgtcgaaatcctcgagggggaagaagg 540
 |||||
 Db 508 agctatcgggttaaagacacctacaggaactactgtcgaaatcctcgagggggaagaagg 567
 |||||
 Qy 541 ggacctggtgtttcacaaagcaatccagaggtacgcgtacgaagtctgtgacattctctcag 600
 |||||
 Db 568 ggacctggtgtttcacaaagcaatccagaggtacgcgtacgaagtctgtgacattctctcag 627
 |||||
 Qy 601 ttttcagaagttgaatgcatacctgcataatgggagagttatcgaggtctcatggatcat 660
 |||||
 Db 628 ttttcagaagttgaatgcatacctgcataatgggagagttatcgaggtctcatggatcat 687
 |||||
 Qy 661 acagaatcaggcaagatttgtcagcgtgggacatcagacacacaccgcgcacaaattc 720
 |||||
 Db 688 acagaatcaggcaagatttgtcagcgtgggacatcagacacacaccgcgcacaaattc 747
 |||||
 Qy 721 ttgcctgaaagatatcccgacaagggttttgataaattattgcgcgaatcccgatggc 780
 |||||
 Db 748 ttgcctgaaagatatcccgacaagggttttgataaattattgcgcgaatcccgatggc 807
 |||||
 Qy 781 cagccgagggccatggtgtatactcttgaccttcacaccgcgtggagtagtctgcaatt 840
 |||||
 Db 808 cagccgagggccatggtgtatactcttgaccttcacaccgcgtggagtagtctgcaatt 867
 |||||
 Qy 841 aaaaatgcgtgcacaaagct 861
 |||||
 Db 868 aaaaatgcgtgcacaaatct 888

RESULT 7

AAQ06088
 ID AAQ06088 standard; cDNA; 2289 BP.

XX AC AAQ06088;

XX DT 31-JAN-1991 (first entry)

XX Tumour cytotoxic factor II gene.

XX

XX PS Claim 1; Fig 6A-D; 19pp; English.

XX CC The nucleotide sequence of human plasminogen-like growth factor

CC PLGF (2172 bp) is used in the stimulation of the growth of endothelial

CC cells or epithelial cells other than hepatocytes. This involves

CC contacting the cells with the PLGF protein or its allelic or splice

CC variants. The growth factor is useful for stimulating wound healing in

CC endothelial or epithelial tissue.

XX SQ Sequence 2172 BP; 674 A; 450 C; 498 G; 550 T; 0 other;

Query Match 49.6%; Score 856.2; DB 19; Length 2172;

Best Local Similarity 99.7%; Pred. No. 1.2e-212;

Matches 858; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atgtgggtgacaaactctcgagccctgctgagcagcatgctctcgcattcctc 60

Db 1 atgtgggtgacaaactctcgagccctgctgagcagcatgctctcgcattcctc 60

QY 61 ctgtcccatcgccatccctctatgcagaggacaaagaaagaaatacaattcat 120

Db 61 ctgtcccatcgccatccctctatgcagaggacaaagaaagaaatacaattcat 120

QY 121 gaattcaaaaaatcgaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

Db 121 gaattcaaaaaatcgaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaagtgaatactgcacacaaatgtgctaataagatgactaggaataaaggactt 240

Db 181 accaaaaagtgaatactgcacacaaatgtgctaataagatgactaggaataaaggactt 240

QY 241 ccattcactgcaaggctttgtttttataagcaagaaacaaatgcctcgtgtccccc 300

Db 241 ccattcactgcaaggctttgtttttataagcaagaaacaaatgcctcgtgtccccc 300

QY 301 ttcaatagcatgcaagtggagtgaaaaagaatttggccatgaatttgacctctatgaa 360

Db 301 ttcaatagcatgcaagtggagtgaaaaagaatttggccatgaatttgacctctatgaa 360

QY 361 acaaaagactacattagaaactgcattgtgtaaaggacgcagctacaaaggaaacagta 420

Db 361 acaaaagactacattagaaactgcattgtgtaaaggacgcagctacaaaggaaacagta 420

QY 421 ttatcactaaagtgggcatcaaatgtagccctgagctgagttccatataccacacaaac 480

Db 421 ttatcactaaagtgggcatcaaatgtagccctgagctgagttccatataccacacaaac 480

QY 481 agctatcggggttaagacactacagaaactactgtcgaatactcgcagggaagaagg 540

Db 481 agctatcggggttaagacactacagaaactactgtcgaatactcgcagggaagaagg 540

QY 541 ggaacctggttttcaagaacatccagaggtacgctacgaagatgtgacattcctcag 600

Db 541 ggaacctggttttcaagaacatccagaggtacgctacgaagatgtgacattcctcag 600

QY 601 tgttcagaagttgaatgcattacacctgcaatgggagagttatcgaggtctctatggaat 660

Db 601 tgttcagaagttgaatgcattacacctgcaatgggagagttatcgaggtctctatggaat 660

QY 661 acagaatcaggcaagatttgcagcgtgggatcatcacaccacacccgcgcacaaattc 720

Db 661 acagaatcaggcaagatttgcagcgtgggatcatcacaccacacccgcgcacaaattc 720

QY 721 ttgctgtaagatacccgacaaggcttttgatgataattattgcgcgaatcccgatg 780

Db 721 ttgctgtaagatacccgacaaggcttttgatgataattattgcgcgaatcccgatg 780

QY 781 cagccgagcccatggtgctatactctgacctctcacaccctggaggtactgtgcaatt 840

Db 781 cagccgagcccatggtgctatactctgacctctcacaccctggaggtactgtgcaatt 840

QY 841 aaacatgcgctgacaaagct 861

Db 841 aaacatgcgctgacaaact 861

RESULT 9

AAQ37308

ID AAQ37308 standard; DNA; 2173 BP.

XX AC AAQ37308;

XX DT 16-JUN-1993 (first entry)

XX DE Encodes haematopoietic stem cell multiplier.

XX KW Bone marrow deficiencies; cancer therapy; tumour; carcinoma;

XX KW bone marrow transplants.

XX FH Key Location/Qualifiers

XX FT CDS 1..2169

XX FT /*tag= a

XX W09303061-A.

XX PD 18-FEB-1993.

XX PF 24-JUL-1992; 92WO-JP00949.

XX PR 26-JUL-1991; 91JP-0187470.

XX PR 26-JUL-1991; 91JP-0187481.

XX PA (TORA) TORAY IND INC.

XX PI Kawano G, Kojima K, Komiyama A, Kubo T, Nakahata T;

XX PI Sano E, Sudot, Tanaka R;

XX DR WPI; 1993-076441/09.

XX DR P-PSDB; AAR32710.

XX PT Haematopoietic stem cell multiplier comprising IL-3 and IL-7 -

XX PT used for treatment and prevention of bone marrow disorders e.g.

XX PS after cancer therapy or bone marrow transplants

XX PS Disclosure; Page 60; 90pp; Japanese.

XX CC This sequence encodes a haematopoietic stem cell multiplier of mol.

XX CC wt. 60,000, and having N-terminal sequence AAR32709.

XX SQ Sequence 2173 BP; 671 A; 452 C; 502 G; 548 T; 0 other;

Query Match 49.6%; Score 856.2; DB 14; Length 2173;

Best Local Similarity 99.7%; Pred. No. 1.2e-212;

Matches 858; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atgtgggtgacaaactctcgcagccctgctgagcagcatgctctcgcattcctc 60

Db 1 atgtgggtgacaaactctcgcagccctgctgagcagcatgctctcgcattcctc 60

QY 61 ctgtcccatcgccatccctctatgcagaggacaaagaaagaaatacaattcat 120

Db 61 ctgtcccatcgccatccctctatgcagaggacaaagaaagaaatacaattcat 120

QY 121 gaattcaaaaaatcgaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

Db 121 gaattcaaaaaatcgaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaagtgaatactgcagaccatgtgctaataagatgactaggaataaaggactt 240

Db 181 accaaaaagtgaatactgcagaccatgtgctaataagatgactaggaataaaggactt 240

QY 241 ccattcactgcaaggctttgtttttataagcaagaaacaaatgcctcgtgtccccc 300

QY 721 ttgctgaagatatcccgacaaaggcttgatgataattattgcccgaatcccgatggc 780
 Db 798 ttgctgaagatatcccgacaaaggcttgatgataattattgcccgaatcccgatggc 857
 QY 781 cagccgagggcattggtctactcttgaccctcacaccgcgtgggagtactgtgcaatt 840
 Db 858 cagccgagggcattggtctactcttgaccctcacaccgcgtgggagtactgtgcaatt 917
 QY 841 aaacatgcgtgacaaagct 861
 Db 918 aaacatgcgtgacaaatct 938

RESULT 11
 AAQ14182
 ID AAQ14182 standard; DNA; 2172 BP.
 AC AAQ14182;
 DT 10-JAN-1992 (first entry)
 DE Plasmidogen-like growth factor gene.
 XX Plasmidogen-like growth factor gene.
 KW Mitogenic activity; PLGF; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH sig_peptide 1..63
 FT /*tag= a
 FT mat_peptide 64..2172
 FT /*tag= b
 PN US7582063-A.
 XX 17-SEP-1991.
 XX 14-SEP-1990; 90US-0582063.
 PR 14-SEP-1990; 90US-0582063.
 PA (USSH) NAT INST OF HEALTH.
 XX Rubin JS, Chan AML, Aaronson SA;
 PI WPI; 1991-317957/43.
 XX P-PSDB; AAR14307.
 DR New plasmidogen-like growth factor protein - having potent
 PT mitogenic activity on melanocytes, epithelial cells and
 PT endothelial cells
 XX Disclosure; Fig 6; 33pp; English.
 PS The DNA sequence is that of a-plasminogen-like growth factor (PGLF)
 CC gene. The DNA encoding PGLF can be obt'd. from a DNA library of M246
 CC human embryonic lung fibroblasts. The sequence may be used in the
 CC determin. of the level of mRNA transcripts encoding the protein
 CC which may be useful in the diagnosis of malignancy.
 XX Sequence 2172 BP; 675 A; 448 C; 499 G; 550 T; 2 other;
 SQ

Query Match 49.4%; Score 851.6; DB 12; Length 2172;
 Best Local Similarity 99.2%; Pred. No. 1.9e-211;
 Matches 854; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 atgtgggtgacaaactctgcagccctgctgagcagcatgctcctcgtcatctctc 60
 Db 1 atgtgggtgacaaactctgcagccctgctgagcagcatgctcctcgtcatctc 60
 QY 61 ctgctcccattcccatccctatgcagaggacaaaggaaagaaataacataatcat 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 ctgctcccattcccatccctatgcagaggacaaaggaaagaaataacataatcat 120
 QY 121 gaattcaaaaaatcagcaaaagactaccctaatcaaaaatagatccagcactgaagataaa 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 gaattcaaaaaatcagcaaaagactaccctaatcaaaaatagatccagcactgaagataaa 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 accaaaaaagtgaatactactgcagaccatgtgctaataagatgtactaggaataaaggact 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 accaaaaaagtgaatactactgcagaccatgtgctaataagatgtactaggaataaaggact 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 ccattcacttcaaggctttttttgtataaagcaaaacaatgcctctggttcccc 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 ccattcacttcaaggctttttttgtataaagcaaaacaatgcctctggttcccc 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 aacaaagactacattagaaaactgcatacttgtaagagcagcagctacaaggaacagta 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 aacaaagactacattagaaaactgcatacttgtaagagcagcagctacaaggaacagta 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 tctatcactaagatggcatcaaatgtcagccctggagtctcatgataccacacgaacac 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 tctatcactaagatggcatcaaatgtcagccctggagtctcatgataccacacgaacac 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 481 agctatcggggttaagacctacagaaaactactgtcgaaaatcctcgaggggaagaagg 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 agctatcggggttaagacctacagaaaactactgtcgaaaatcctcgaggggaagaagg 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 541 ggaacctggttttcacaaagcaatccagaggttacgctacgaaagtctgtgacattctcag 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 ggaacctggttttcacaaagcaatccagaggttacgctacgaaagtctgtgacattctcag 600
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 601 tgttcagaagttgaatgcatacctgcattggggagagttatcgaggtctctatggatcat 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 tgttcagaagttgaatgcatacctgcattggggagagttatcgaggtctctatggatcat 660
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 661 acagaatcaggcaagatttgcagcgtgggatcatcacaccacacccgacacaaatc 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 661 acagaatcaggcaagatttgcagcgtgggatcatcacaccacacccgacacaaatc 720
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 721 ttgctgaagatatcccgacaaaggcttgatgataattattgcccgaatcccgatggc 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 721 ttgctgaagatatcccgacaaaggcttgatgataattattgcccgaatcccgatggc 780
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 781 cagccgagggcattggtctactcttgaccctcacaccgcgtgggagtactgtgcaatt 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 781 cagccgagggcattggtctactcttgaccctcacaccgcgtgggagtactgtgcaatt 840
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 841 aaacatgcgtgacaaagct 861
 |||||||||||||||||| ||
 Db 841 aaacatgcgtgacaaatct 861

RESULT 12

AAQ47832

ID AAQ47832 standard; cDNA to mRNA; 2187 BP.

XX AC AAQ47832;

XX AC AAQ47832;

XX DT 14-MAR-1994 (first entry)

XX DE Competative inhibitor of HGF.

XX DE HGF; hepatocyte growth factor; transformation; antagonist;

XX KW liver disease; ss.

XX KW Homo sapiens.

XX OS

XX XX

XX Key Location/Qualifiers

XX FH 1..2187

XX FT CDS /*tag= a

XX FT

DR P-PSDB; AAR40863.

XX New protein is competitive inhibitor of hepatocyte growth factor
PT - used in animal model of liver disease

XX

PS Claim 6; Page 11-14; 20pp; Japanese.

XX

CC The gene encodes an inhibitor of HGF, which has a molecular weight
of ca. 76000 to 92000 as determined by SDS-PAGE. Human modified
CC (NP) HGF-A and -B fragments were prepared from human HGF cDNA.
CC This was amplified by PCR to give a human NP-HGF DNA fragment that
CC was transformed into E.coli. (see AAQ47833) for a related sequence.

XX

SQ Sequence 2187 BP; 667 A; 454 C; 500 G; 557 T; 9 other;

Query Match 48.3%; Score 832.8; DB 14; Length 2187;

Best Local Similarity 98.1%; Pred. No. 1.5e-206;

Matches 859; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 atgtgggtgacaaaactcctgccagccctgtctgtcagcagatgtcctcctgcattcctc 60

Db 1 atgtgggtgacaaaactcctgccagccctgtctgtcagcagatgtcctcctgcattcctc 60

QY 61 ctgtcccccattgcccattccctatgcagaggggacaaaggaaagaaatacaattcat 120

Db 61 ctgtcccccattgcccattccctatgcagaggggacaaaggaaagaaatacaattcat 120

QY 121 gaattcaaaaaatcagaaagactaccctaatcaaaatagatccagcactgaagataaaa 180

Db 121 gaattcaaaaaatcagaaagactaccctaatcaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaagtgaattactgcagaccatgtgctaatagtactagataaaaggactt 240

Db 181 accaaaaagtgaattactgcagaccatgtgctaatagtactagataaaaggactt 240

QY 241 ccattcactgcagggtttttgttttgcataaagcagaaacaaatgcctgtgtccccc 300

Db 241 ccattcactgcagggtttttgttttgcataaagcagaaacaaatgcctgtgtgtccccc 300

QY 301 ttcattagcgtcgaagtgcagtgagtgaaagaaatgttgccatgaattgacctctatgaa 360

Db 301 ttcattagcgtcgaagtgcagtgagtgaaagaaatgttgccatgaattgacctctatgaa 360

QY 361 acaaaagactacattagaaactgcattctgtgtaaaagcagcagctacaagggaacagta 420

Db 361 acaaaagactacattagaaactgcattctgtgtaaaagcagcagctacaagggaacagta 420

QY 421 tctatcactaagagtgccatcaaatgtcagccctggagttccatgatcacacgaacac 480

Db 421 tctatcactaagagtgccatcaaatgtcagccctggagttccatgatcacacgaacac 480

QY 481 -----agctatcggggttaaagacctacagaaaactactgtcgaattcct 525

Db 481 agcttttgccttcgagctatcggtgggcaagacctacagaaaactactgtcgaattcct 540

QY 526 cgaggggaaaggggagccctgtgtttcacaaagcaatccagaggtacgtacgaagtc 585

Db 541 cgaggggaaaggggagccctgtgtttcacaaagcaatccagaggtacgtacgaagtc 600

QY 586 tctgacattcctcagtttccagagttgaatgcagcctgcaatggggagagttatcga 645

Db 601 tctgacattcctcagtttccagagttgaatgcagcctgcaatggggagagttatcga 660

QY 646 ggtctcatgatcatcacagatcaggcaagatttgcagcgtgggcatcatcagaaccca 705

Db 661 ggtctcatgatcatcacagatcaggcaagatttgcagcgtgggcatcatcagaaccca 720

QY 706 caccggcacaaattcttgctgaaagatatcccgacaagggcttgcagataattatgc 765

Db 721 caccggcacaaattcttgctgaaagatatcccgacaagggcttgcagataattatgc 780

QY 766 cgcaatccccgatggccagccgagggccatggtgtctatctcttgacctcacaccgcgtg 825

Db 781 cgcatacccgatggccagccgagccatggtgtctatactctgacctcacaccgcgtg 840

QY 826 gagtactgtcaattaaacacatgcgtgcacaaagt 861

Db 841 gagtactgtcaattaaacacatgcgtgcacaaagt 876

RESULT 14

AAQ45702

ID AAQ45702 standard; DNA; 2187 BP.

XX AAQ45702;

AC AAQ45702;

XX 15-NOV-1993 (first entry)

XX Vascular endothelial cell growth enhancer DNA.

XX Enhance; growth; vascular endothelial cell; human; tumour; cell line;

XX HUOCA-II; HUOCA-III; blood vessel; wounds; burns; decubitus;

XX post-operative tissue damage; drug; cardiac angiopathy; ss.

XX Homo sapiens.

XX EP550296-A.

XX 07-JUL-1993.

XX 27-NOV-1992; 92EP-0403199.

XX 28-NOV-1991; 91JP-0337999.

XX (TERU) TERUMO CORP.

XX Adachi M, Harada K, Hirahara I, Sudo T;

XX WPI; 1993-215669/27.

XX P-PSDB; AAR42062.

XX Vascular endothelial cell growth factor protein - used for

XX promoting angiogenesis in the treatment of cardiac angiopathy,

XX wounds, burn injuries, postoperative tissue damage etc.

XX Claim 5; Page 33-35; 44pp; English.

XX This sequence encodes a single chain protein which selectively

XX enhances the growth of vascular endothelial cells. This protein

XX was produced by the human tumour cell line HUOCA-II or HUOCA-III.

XX This protein enhances the formation of new blood vessels and may be

XX used to enhance healing of wounds, burns injuries, decubitus or

XX post-operative tissue damage. It may also be used as a drug for

XX cardiac angiopathy.

XX Sequence 2187 BP; 674 A; 455 C; 501 G; 557 T; 0 other;

SQ

Query Match

Best Local Similarity 98.1%; Pred. No. 1.5e-206;

Matches 859; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 atgtgggtgacaaaactcctgccagccctgtctgtcagcagatgtcctcctgcattcctc 60

Db 1 atgtgggtgacaaaactcctgccagccctgtctgtcagcagatgtcctcctgcattcctc 60

QY 61 ctgtccccctcgcctatccctatcagagggacaaagaaagaaatacaattcat 120

Db 61 ctgtccccctcgcctatccctatcagagggacaaagaaagaaatacaattcat 120

QY 121 gaattcaaaaaatcagcaaaagactaccctaatcaaaatagatccagcactgaagataaaa 180

Db 121 gaattcaaaaaatcagcaaaagactaccctaatcaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaagtgaattactgcagaccatgtgctaatagtactaggataaaaggactt 240


```
QY 586 tgtgacattcctcagtggtcagaagtgaatgcatgacctgcaatgggagagattatcga 645
Db 628 tgtgacattcctcagtggtcagaagtgaatgcatgacctgcaatgggagagattatcga 687
QY 646 ggtctcatggtatcacacagaatcaggcaagatttgtcagcgtggggtatcatcagacacca 705
Db 688 ggtctcatggtatcacacagaatcaggcaagatttgtcagcgtggggtatcatcagacacca 747
QY 706 caccggcacaaattcttgctgaaagatatcccgacaagggttttgatgataattattgc 765
Db 748 caccggcacaaattcttgctgaaagatatcccgacaagggttttgatgataattattgc 807
QY 766 cgcaatcccgatggccagccgagccatggtgctatactcttgacccctcacacccgcctgg 825
Db 808 cgcaatcccgatggccagccgagccatggtgctatactcttgacccctcacacccgcctgg 867
QY 826 gagtactgtgcaattaaaacatgcgctgacaaagct 861
Db 868 gagtactgtgcaattaaaacatgcgctgacaaatct 903
```

Search completed: September 18, 2001, 18:33:25
Job time: 7482 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:21:47 ; Search time 97.89 Seconds
(without alignments)
3336.007 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 agtgggggacccaactctct.....accacaccaccaccactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	857.8	49.7	2289	1	US-07-838-410-2
2	856.2	49.6	2172	4	US-08-030-410-2
3	845.8	49.0	2286	1	US-08-290-937B-4
4	832.8	48.3	2184	1	US-07-815-333A-1
5	159	9.2	2021	1	US-07-882-925A-3
6	159	9.2	2021	1	US-08-184-012C-3
7	159	9.2	2216	2	US-08-666-082B-2
8	159	9.2	2219	1	US-07-882-925A-1
9	159	9.2	2219	1	US-08-184-012C-1
10	159	9.2	2232	1	US-08-334-177-1
11	159	9.2	2232	5	PCT-US95-13830-1
12	159	9.2	2262	1	US-07-882-925A-7
13	159	9.2	2262	1	US-08-184-012C-7
14	157.4	9.1	2219	1	US-07-882-925A-2
15	157.4	9.1	2219	1	US-08-184-012C-2
16	146	8.5	2188	1	US-07-882-925A-4
17	146	8.5	2188	1	US-08-184-012C-4
18	136.4	7.9	1284	3	US-08-985-526-24
19	136.2	7.9	645	3	US-08-985-526-22
20	136	7.9	2497	1	US-08-643-219-12
21	136	7.9	2497	2	US-09-131-995-12
22	136	7.9	2497	2	US-08-832-087B-12
23	136	7.9	2497	3	US-08-851-350-12
24	136	7.9	2497	4	US-09-132-154-12
25	136	7.9	2679	6	5200340-7
26	134.4	7.8	2753	1	US-07-854-603-1
27	134	7.8	1134	4	US-09-206-059-29

28	128.4	7.4	2296	1	US-07-750-080A-18	Sequence 18, Appl
29	128.4	7.4	2296	3	US-08-651-472-18	Sequence 18, Appl
30	55.4	3.2	3398	5	PCT-US95-08493-12	Sequence 12, Appl
31	50.8	2.9	6751	1	US-07-882-925A-5	Sequence 5, Appl
32	50.8	2.9	6751	1	US-08-184-012C-5	Sequence 5, Appl
33	50	2.9	6100	1	US-07-882-925A-6	Sequence 6, Appl
34	50	2.9	6100	1	US-08-184-012C-6	Sequence 6, Appl
35	47.8	2.8	7218	1	US-08-232-463-14	Sequence 14, Appl
36	46.8	2.7	175	1	US-08-235-838-3	Sequence 3, Appl
37	46.8	2.7	175	2	US-08-465-473B-3	Sequence 3, Appl
38	46.8	2.7	720	2	US-08-800-198-7	Sequence 7, Appl
39	46.8	2.7	720	3	US-09-296-595-7	Sequence 7, Appl
40	46.8	2.7	748	1	US-08-235-838-4	Sequence 4, Appl
41	46.8	2.7	748	1	US-08-235-838-10	Sequence 10, Appl
42	46.8	2.7	748	2	US-08-465-473B-4	Sequence 4, Appl
43	46.8	2.7	748	2	US-08-465-473B-10	Sequence 10, Appl
44	46.8	2.7	2012	1	US-08-235-838-13	Sequence 13, Appl
45	46.8	2.7	2012	1	US-08-235-838-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-07-838-410-2
; Sequence 2, Application US/07838410
; Patent No. 5328836
; GENERAL INFORMATION:
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: HIGASHIO, KANJI
; APPLICANT: NAGAO, MASAYA
; APPLICANT: OOGAKI, FUMIKO
; APPLICANT: TAKAOKA, HIROAKI
; APPLICANT: TSUDA, EISUKE
; TITLE OF INVENTION: PLASMIDS CONTAINING DNA ENCODING THE
; TITLE OF INVENTION: AMINO ACID SEQUENCE OF TCF-II, TRANSFORMED CELLS WITH THE
; TITLE OF INVENTION: PLASMIDS, AND PRODUCTION METHOD OF BIOLOGICALLY ACTIVE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD.
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/838,410
; FILING DATE: 19920311
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP91/00942
; FILING DATE: 15-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185852-1990
; FILING DATE: 13-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: WHITE, JOHN L.
; REGISTRATION NUMBER: 17,746
; REFERENCE/DOCKET NUMBER: WAK 110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; TELEX: 64191
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2289 base pairs
; TYPE: NUCLEIC ACID


```

SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,925A
FILING DATE: 19920514
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Gregory
REGISTRATION NUMBER: 29,945
REFERENCE/DOCKET NUMBER: CMC 57
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 421-2324
TELEFAX: (513) 421-7269
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2219 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #33
POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/p33F
FEATURE:
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: Includes five
OTHER INFORMATION: level; one of w
OTHER INFORMATION: 619). Sequence
OTHER INFORMATION: polymorphic ami
OTHER INFORMATION:
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1:
US-07-882-925A-1

```

Query Match 9.2%; Score 159; DB 1; Length 2219;
Best Local Similarity 59.1%; Pred. No. 3.3e-33;
Matches 295; Conservative 0; Mismatches 195; Indels 9; Gaps 1;

[illegible]

```

Db      696  AACTATTGCGCGAATCTGACGGCTCGGAGCGGCATGGTGTACTACGAGATCCGCAG 755
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy      817  acccgctgggagtactgtg 835
      || ||||| ||||| |||||
Db      756  ATCGAGCGAGAGTTCTCTG 774
      .

RESULT          9
US-08-184-012C-1
; Sequence 1, Application US/08184012C
; Patent No. 5606029
; GENERAL INFORMATION:
; APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: protein
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gregory Lunn
; STREET: Wood, Herron & Evans, 2700 Carew Tower
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5.2
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,012C
; FILING DATE: 1/18/94

```

```

; ; OTHER INFORMATION: level; one of which results in an amino acid substitution (nu
; ; OTHER INFORMATION: 619). Sequence ID NO:2: contains the identical sequence with
; ; OTHER INFORMATION: polymorphic amino acid.
; ; PUBLICATION INFORMATION:
; ; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2219
US-07-882-925A-1

```

TELEFAX: (313) 421-7209
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2219 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #33
POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/D3F15S2
FEATURE:
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: Includes five polymorphisms at the nucleotide
level; one of which results in an amino acid substitution
OTHER INFORMATION: 619). Sequence ID NO:2: contains the identical sequence w
Patent No. 5606029
OTHER INFORMATION: polymorphic amino acid.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2219
US-08-184-012C-1

```
Query Match          9.2%; Score 159; DB 1; Length 2219;
Best Local Similarity 59.1%; Pred. NO. 3.3e-33;
Matches 295; Conservative 0; Mismatches 195; Indels 9; Gaps 1;
```

```
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0912PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2232 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
PCT-US95-13830-1

Query Match          9.2%; Score 159; DB 5; Length 2232;
Best Local Similarity 59.1%; Pred. No. 3.3e-33;
Matches 295; Conservative 0; Mismatches 195; Indels 9; Gaps 1;

Qy 346 tttagacctatgaaacaaagactacattagaacatgcatcattggttaaaggacgcagc 405
Db 302 TGTGACCTCTTCCAGAGAAAGACTACGTACGGACCTGCATCATGAACAATGGGGTTGGG 361
Qy 406 tacaaggaaacagtgatctatcattcaagtgatggtcacaataatgcagccctggagttccatg 465
Db 362 TACCGGGCCACATGGCCAGCCGTGGTGGCTGCCCTGCCAGCTTGGAGCCACAG 421
Qy 466 ataccacacgaacacagcta-----tcggggtaaaagacctacagggaaactactgt 516
Db 422 TTCCCGAATGATACAAAGTACACGCCCACTCTCCGGAATGCCCTGGAAGAGAACTTCGC 481
Qy 517 cgaatcctcgagggaagaagggggaccctgtgttttcaacaagaatccagaggtacgc 576
Db 482 CGTAACCTGTATGCGCGACCCCGGAGGTCCTTGGTGTCTACACACAGACCCCTGCTGGCG 541
Qy 577 tacgaagctgtgacattcctcagttcagaaagtgatgaatgacatgacacctgcaatgggag 636
Db 542 TTCAGAGCTGCGGCATCAAAATCTGCGGGAGCGCGGTGTCTGTGTGCAATGGCGAG 601
Qy 637 agttcagagtgctcatggtatcatcacagaatcaggcaagatttgcagcgctgggacat 696
Db 602 GAATACCGCGCGCGGTAGACCGCACGAGTCAAGGCGCGAGTGCACGCGTGGGATCTT 661
Qy 697 cagacacacacccggcacaattcttgcctgaaagatatcccgacaagaaggcttggatgat 756
Db 662 CAGCACCCGCGCACACCCCTTCGAGCGGGCAAGTTCCCTCGACCAAGGCTCGACGAC 721
Qy 757 aattatgctgcgaatcccgatggccagccagcgccatgggtgctatactcttgaccctcac 816
Db 722 AACTATGCGGAAATCTGACGGCTCCGAGCGGCCATGGTGTCTACACTACCGATCCGCGAG 781
Qy 817 acccgctgggagtagtctgtg 835
Db 782 ATCAGCGAGAGTTCTGTG 800

RESULT 12
US-07-882-925A-7
; Sequence 7, Application US/07882925A
; Patent No. 5315000
; GENERAL INFORMATION:
; -APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; -ADDRESSEE: Gregory Lunn
; STREET: Wood, Herron & Evans, 2700 Carew Tower
; CITY: Cincinnati
```

```
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.3
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/882,925A
; FILING DATE: 19920514
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Gregory
; REGISTRATION NUMBER: 29,945
; REFERENCE/DOCKET NUMBER: CMC 57
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 241-2324
; TELEFAX: (513) 421-7269
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2262 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; DESCRIPTION: Identical to sequence ID NO: 1: with 5' and 3'
; DESCRIPTION: adaptors added to make a full-length cDNA
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; DEVELOPMENTAL STAGE: fetal
; TISSUE TYPE: liver
; IMMEDIATE SOURCE:
; LIBRARY: cDNA
; CLONE: #33 including 5' and 3' adaptors
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 2262
US-07-882-925A-7

Query Match          9.2%; Score 159; DB 1; Length 2262;
Best Local Similarity 59.1%; Pred. No. 3.3e-33;
Matches 295; Conservative 0; Mismatches 195; Indels 9; Gaps 1;

Qy 346 tttagacctatgaaacaaagactacattagaacatgcatcattggttaaaggacgcagc 405
Db 307 TGTGACCTCTTCCAGAGAAAGACTACGTACGGACCTGCATCATGAACAATGGGGTTGGG 366
Qy 406 tacaaggaaacagtgatctatcattcaagtgatggtcacaataatgcagccctggagttccatg 465
Db 367 TACCGGGCCACATGGCCAGCCGTGGTGGCTGCCCTGCCAGCTTGGAGCCACAG 426
Qy 466 ataccacacgaacacagcta-----tcggggtaaaagacctacagggaaactactgt 516
Db 427 TTCCCGAATGATACAAAGTACACGCCCACTCTCCGGAATGCCCTGGAAGAGAACTTCGC 486
Qy 517 cgaatcctcgagggaagaagggggaccctgtgtttcacaagaatccagaggtacgc 576
Db 487 CGTAACCTGTATGCGCGACCCCGGAGGTCCTTGGTGTCTACACACAGACCCCTGCTGTGCGC 546
Qy 577 tacgaagctgtgacattcctcagttgtaagttgaatgcatgacctggagggag 636
Db 547 TTCCAGAGCTGCGGCATCAAAATCTGCCGGAGCGCGGTGTCTGTGTGCAATGGCGAG 606
Qy 637 agttcagagtgctcatggtatcatcacagaatcaggcaagaatttgcagcgctgggacat 696
Db 607 GAATACCGCGCGCGGTAGACCGCACGAGTCAAGGCGCGAGTGCACGCGCTGGGATCTT 666
Qy 697 cagacacacacccggcacaattcttgcctgaaagatatcccgacaagaaggcttggatgat 756
Db 667 CAGCACCCGCGCACACCCCTTCGAGCGGCCGCAAGTTCTCTCGACCAAGGCTCTGGAGCAC 726
```

DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: CDNA
CLONE: #33
POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/D3F15S2
FEATURE:
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: Includes five polymorphisms at the nucleotide
OTHER INFORMATION: level; one of which results in an amino acid substitution (nu
OTHER INFORMATION: 619). Sequence ID NO:1: contains the identical sequence with
OTHER INFORMATION: polymorphic amino acid.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2219
US-07-882-925A-2

Query Match 9.1%; Score 157.4; DB 1; Length 2219;
Best Local Similarity 58.9%; Pred. No. 8.9e-33;
Matches 294; Conservative 0; Mismatches 196; Indels 9; Gaps 1;
QY 346 ttgacctctatgaaacaaagactacattagaactgcatcattgttaaagagcgcagc 405
DB 276 TGTGACCTCTCCAGAGAAAGACTAGTACGGACCTGTCATCATGAACAATGGGTGGG 335
QY 406 tacaagggaacagtgatctatcactaagaagtggaacatcaaatgtcagccctgaggttccatg 465
DB 336 TACCGGGGACCATGGCCACGACCGTGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAG 395
QY 466 ataccacacgaacacagcta-----tcggggtaagaacacctacagggaaactactgt 516
DB 396 TTCCCGAATGATCAACAAGTACACGCCCACTCTCCGGAATGGCCTGGAGAGAACTTCTGC 455
QY 517 cgaatctcaggggaagaagggggaccctgtgtttcacaagaatccagaggtacgc 576
DB 456 CGTAACCTGTAGTGGCGAGCCCGGAGGCTCTGGTGTCTACACACAGACCTCTGGTGGC 515
QY 577 tacaagtgctgacattctcagttcagaaagttgaatgcatgacctgcgaatggggag 636
DB 516 TTCCAGAGCTGCGGCATCAAAATCTCCGCGGAGGCGCGTGTGTGTGTAATGGCGAG 575
QY 637 agttatcagaggtctcattggtatcacaagaatcaggcaagattgtcagcgtggatcat 696
DB 576 GAATACCGCGCGCGGTAGACCGCAGCGAGTCCAGGGCGGAGTTCACGCGTGGGATCTT 635
QY 697 cagacacacacgggcacaaattcttcctgaaagatatcccgacaagggtcttgatgat 756
DB 636 CAGACCGCGACACGACCCCTTCGACCGGGCAAGTCTCTGACCAAGGCTCGACGAC 695
QY 757 aattatgcgcgaatcccgatggccagcgagggccatggtgtctatactcttgaccctcac 816
DB 696 AACTATTCCCGGAATCTGACGGCTCCGAGCGGCATGGTGTCTACACTACGGATCCGCGAG 755
QY 817 acccgctggaggtactgtg 835
DB 756 ATCAGCGGAGAGTTCTGTG 774

RESULT 15

US-08-184-012C-2
Sequence 2, Application US/08184012C
Patent No. 5606029
GENERAL INFORMATION:
APPLICANT: Degen, Sandra J. F.
TITLE OF INVENTION: Gene for a growth factor and its cDNA and
NUMBER OF INVENTIONS: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Gregory Lunn
STREET: Wood, Herron & Evans, 2700 Carew Tower
CITY: Cincinnati
STATE: Ohio

COUNTRY: USA
ZIP: 45202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 800 KB
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.2
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,012C
FILING DATE: 1/18/94
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Gregory
REGISTRATION NUMBER: 29,945
REFERENCE/DOCKET NUMBER: CMC 57
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 241-2324
TELEFAX: (513) 421-7269
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2219 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: CDNA
CLONE: #33
POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/D3F15S2
FEATURE:
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: Includes five polymorphisms at the nucleotide
OTHER INFORMATION: level; one of which results in an amino acid substitution
OTHER INFORMATION: 619). Sequence ID NO:1: contains the identical sequence w
Patent No. 5606029
OTHER INFORMATION: polymorphic amino acid.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2219
US-08-184-012C-2

Query Match 9.1%; Score 157.4; DB 1; Length 2219;
Best Local Similarity 58.9%; Pred. No. 8.9e-33;
Matches 294; Conservative 0; Mismatches 196; Indels 9; Gaps 1;
QY 346 ttgacctctatgaaacaaagactacattagaactgcatcattgttaaagagcgcagc 405
DB 276 TGTGACCTCTCCAGAGAAAGACTAGTACGGACCTGTCATCATGAACAATGGGTGGG 335
QY 406 tacaagggaacagtgatctatcactaagaagtggaacatcaaatgtcagccctgaggttccatg 465
DB 336 TACCGGGGACCATGGCCACGACCGTGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAG 395
QY 466 ataccacacgaacacagcta-----tcggggtaagaacacctacagggaaactactgt 516
DB 396 TTCCCGAATGATCAACAAGTACACGCCCACTCTCCGGAATGGCCTGGAGAGAACTTCTGC 455
QY 517 cgaatctcaggggaagaagggggaccctgtgtttcacaagaatccagaggtacgc 576
DB 456 CGTAACCTGTAGTGGCGAGCCCGGAGGCTCTGGTGTCTACACACAGACCTCTGGTGGC 515
QY 577 tacaagtgctgacattctcagttcagaaagttgaatgcatgacctgcgaatggggag 636
DB 516 TTCCAGAGCTGCGGCATCAAAATCTCCGCGGAGGCGCGTGTGTGTGTAATGGCGAG 575
QY 637 agttatcagaggtctcattggtatcacaagaatcaggcaagattgtcagcgtggatcat 696

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:21:42 ; Search time 2582.57 Seconds
(without alignments)
6313.932 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtgggtaccacaaactcct.....accacacacacacacactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estp11:*
73: em_estp12:*
74: em_estp13:*
75: em_estp14:*
76: em_estp15:*
77: em_estp16:*
78: em_estp17:*
79: em_estp18:*
80: em_estp19:*
81: em_estp110:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

SUMMARIES

Result No.	Query	Score	Length	DB	ID	Description
1	AL546558	47.6	975	106	AL546558	AL546558
2	AL549164	46.0	919	106	AL549164	AL549164
3	AL550803	38.0	872	106	AL550803	AL550803
4	AL546883	35.5	828	106	AL546883	AL546883
5	AA037738	25.3	506	1	AA037738	AA037738
6	AL546856	43.2	25.0	844	AL546856	AL546856
7	AL573472	22.6	899	106	AL573472	AL573472
8	AV759905	15.7	412	110	AV759905	AV759905
9	AA037786	9.6	484	1	AA037786	AA037786
10	AV707491	9.5	415	32	AV707491	AV707491
11	BE244598	9.3	321	166	BE244598	BE244598
12	AW175268	7.9	661	112	AW175268	AW175268
13	AQ381428	7.5	753	227	AQ381428	AQ381428
14	BF383652	7.7	766	148	BF383652	BF383652
15	AI317783	7.6	778	18	AI317783	AI317783
16	BF079866	7.6	565	144	BF079866	BF079866
17	BE755132	7.4	518	139	BE755132	BE755132
18	AI317745	7.4	768	18	AI317745	AI317745
19	AI118965	7.3	837	16	AI118965	AI118965
20	AI530175	7.2	844	21	AI530175	AI530175
21	AQ707814	7.1	470	232	AQ707814	AQ707814
22	BE326846	7.1	470	166	BE326846	BE326846
23	AI317794	7.0	780	18	AI317794	AI317794
24	AL531543	6.9	914	106	AL531543	AL531543
25	AL557971	6.8	909	106	AL557971	AL557971
26	AI226536	6.7	980	17	AI226536	AI226536
27	AI121047	6.5	788	16	AI121047	AI121047
28	AI114957	6.5	819	16	AI114957	AI114957
29	AI746861	6.5	817	24	AI746861	AI746861
30	BE576686	6.4	265	137	BE576686	BE576686
31	AI119053	6.4	773	16	AI119053	AI119053
32	AW106317	6.4	897	111	AW106317	AW106317
33	AI663558	6.3	918	23	AI663558	AI663558
34	BF235733	6.2	752	146	BF235733	BF235733
35	BF234001	6.2	780	146	BF234001	BF234001
36	AI118969	6.2	842	16	AI118969	AI118969
37	AI530076	6.1	910	21	AI530076	AI530076
38	AL558014	5.9	889	106	AL558014	AL558014
39	AI746972	5.9	794	24	AI746972	AI746972
40	AI119056	5.7	741	16	AI119056	AI119056
41	BF385803	5.6	843	148	BF385803	BF385803
42	AZ694675	5.4	243	248	AZ694675	AZ694675
43	AA260053	5.4	403	4	AA260053	AA260053
44	BE234370	5.4	574	165	BE234370	BE234370
45	AI958441	5.4	594	104	AI958441	AI958441

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCSSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	AL546558	AL546558 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YH12 5 prime, mRNA sequence.	AL546558	EST	16-FEB-2001	human	Homo sapiens	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	1 (bases 1 to 975)	Li, W. B., Gruber, C., Jessee, J. and Polayes, D.	Full-length cDNA libraries and normalization	Unpublished (2001)
											Contact: Genoscope	

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1 . 975

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI030YH12"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT 311 a 228 c 215 g 220 t 1 others
ORIGIN

Query Match 47.6%; Score 820.8; DB 106; Length 975;

Best Local Similarity 97.9%; Pred. No. 5.6e-211; Indels 16; Gaps 2;

Matches 858; Conservative 0; Mismatches 2;

QY 1 atgtgggtgacaaacctctgcagccctgctgctgaggggcaagaaagaaagaaatacaatcat 120

Db 6 atgtgggtgacaaacctctgcagccctgctgctgaggggcaagaaagaaagaaatacaatcat 125

QY 61 ctgtctccctcgcctccctcctatgagggggcaagaaagaaagaaatacaatcat 180

Db 66 ctgtctccctcgcctccctcctatgagggggcaagaaagaaagaaatacaatcat 185

QY 121 gaattcaaaaaatcagcaagactaccctaaatacaatgaggggcaagaaagaaatacaatcat 180

Db 126 gaattcaaaaaatcagcaagactaccctaaatacaatgaggggcaagaaagaaatacaatcat 185

QY 181 accaaagagtgaaactactcagaccactgtgctaatagatgtaggaggaataagagactt 240

Db 186 accaaagagtgaaactactcagaccactgtgctaatagatgtaggaggaataagagactt 245

QY 241 ccattcacttgcaagggctttgtttttgataagcaagaaacaaatgcctctggtcccc 300

Db 246 ccattcacttgcaagggctttgtttttgataagcaagaaacaaatgcctctggtcccc 305

QY 301 ttcaatagcatgcaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 360

Db 306 ttcaatagcatgcaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 365

QY 361 aacaagactatagaaactgcattgtaagggcagcagctacaggaagacagta 420

Db 366 aacaagactatagaaactgcattgtaagggcagcagctacaggaagacagta 425

QY 421 tctatcactaaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 480

Db 426 tctatcactaaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 485

QY 481 -----agctatcggggtaagacactcaggaagaaactactctgcaaatcct 525

Db 486 agctatcggggtaagacactcaggaagaaactactctgcaaatcct 545

QY 526 cgagggggaagagggggacccctggtgttttcaagcaagcaatccagaggtacgtcagaagtc 585

Db 546 cgagggggaagagggggacccctggtgttttcaagcaagcaatccagaggtacgtcagaagtc 605

QY 586 tctgacattcctcagtggttcagaagtgagtgagtgagtgagtgagtgagtgagtgagtgag 645

Db 606 tctgacattcctcagtggttcagaagtgagtgagtgagtgagtgagtgagtgagtgagtgag 665

QY 646 ggtctcatggtatcatagacaatcaggcaagatttgtcagcgctgggagatcatcagacacca 705

```

/clone="CS0DI056YB22"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes=Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

```

BASE COUNT      274 a   212 c   110 g   266 t   10 others
ORIGIN

```

```

Query Match      38.0%; Score 656.2; DB 106; Length 872;
Best Local Similarity 89.3%; Pred. No. 1.7e-166;
Matches 735; Conservative 9; Mismatches 62; Indels 17; Gaps 3;

```

```

Qy 1 atgtgggtgacaaactcctgagccctgctgctgagcagatgctcctgcatctctc 60
Db 51 ATTTTATTTTACCAACTCTTCCATCCCTTCTTCTTCATCATTTCCCTTCATCTCTC 110
Qy 61 ctgctccccatgccatccccctgctgctgagcagcagcagcagcagcagcagcagc 120
Db 111 CTTCTCCCATCGCATCCCTATTCATATTGACAAATTAATAATAATAATAATAATTCAT 170
Qy 121 gaattcaaaaaatcagcaagaactaccctaatcaaaatagatccagcagcagcagcagc 180
Db 171 TAATCAAAAAATCATCAAAATACCTTACCTTAATCAAAATATATCATCATTAATAATAA 230
Qy 181 accaaaaagtgaactcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 240
Db 231 ACCAAAAATTTAATCTGATACCAATTTCTTAATATATTTACTATTAATAATAATTA 290
Qy 241 ccattcacttgcagggcttttttttttttttttttttttttttttttttttttttttt 300
Db 291 CCAATTCATCTTCAAKKCTTTTTTTTTTTTATATAATCAAGAAACAAATCTCTCTTCCC 350
Qy 301 tcaatagatgtcaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 360
Db 351 TTCAATAGCATTTCAAGTTGAGTTAAATAAATAAATTTTGCCATTAAATTTACCTCTAT 410
Qy 361 acaaaagactacattagaaactgcatctgtgtaagagcagcagcagcagcagcagcagc 420
Db 411 ACANAGACTACATTAKAACTTCATCATTTGTTAAAGCAGCAGCTACAGGGAACAGTA 470
Qy 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacacagcagc 480
Db 471 TCTATCATAAGAGTGCATCAAAATGTCAGCCCTGGAGTTCCATGATACACACTAACAC 530
Qy 481 -----agctatcggtgtaagacactacagagaaactactgtcgaatccct 525
Db 531 ATCTTTTCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 590
Qy 526 cgaggggaagaggggagccctgtgtttcacagcaatccagaggtacgctcagcagcagc 585
Db 591 CTAYGTGAATAMKGTGA-CTGTGTGTTTCCAAAGCAATCCAGAGGTACCTACGAAGTC 649
Qy 586 tgt-gacattcctcagtttcagaagtgtgcatgaccttgcgaatgggagagttatcg 644
Db 650 TGTGGACATTCCTCAGTGTTCAGAAGTTGAATGCATGACCTGCAATGGGAGAGTTATCG 709
Qy 645 aggtcctatgatacatagaaatcagcagagatttgcagcgtggagatcatcagacacc 704
Db 710 AGGTCTCATGGATCATACAGAATCAGCAAGATTTGTGTCAGCGTGGGATCATCAGACACC 769
Qy 705 acaccggcaaaattcttgcctaaagatatccgacaagggctttgatataatttg 764
Db 770 ACACCGGCACAAATCTTGCTGCTGAAAGATATCCGACAAGGGCTTTTCATGATAATATTG 829

```

```

Qy 765 ccgcaatcccatgagccagccagcagcagcagcagcagcagcagcagcagcagcagc 807
Db 830 CGCAATCCCATGATGCCAGCGAGGCGCCATGCTGCTATCTCTT 872

RESULT 4
AL546883 AL546883 828 bp mRNA EST 16-FEB-2001
LOCUS AL546883 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YD16 5
DEFINITION prime, mRNA sequence.
ACCESSION AL546883
VERSION AL546883.1 GI:12880433
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..828
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YD16"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes=Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      290 a   183 c   157 g   196 t   2 others
ORIGIN

Query Match      35.5%; Score 612; DB 106; Length 828;
Best Local Similarity 99.4%; Pred. No. 1.4e-154;
Matches 612; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgtgggtgacaaactcctgagccctgctgctgagcagcagcagcagcagcagcagc 60
Db 15 ATGTGGGTGACCAACTCTCTGCCAGCCCTGCTGTGTCAGCATGTCTCTGTCATCTCTC 74
Qy 61 ctgctccccatgcccctcctctatgagaggagcagcagcagcagcagcagcagcagcagc 120
Db 75 CTGCTCCCATCGCATCCCTATGTCAGAGGACAAAGGAGAAAGAAATACAAATTCAT 134
Qy 121 gaattcaaaaaatcagcaagaactaccctaatcaaaatagatccagcagcagcagcagc 180
Db 135 GAATTCAAAAATCAGCAAGACTACCCCTAATCAAAATAGATCCAGCATGAGATAAAA 194
Qy 181 accaaaaagtgaatactgagaccacaaatgtgctaataatgactagggaataaaggactt 240
Db 195 ACCAAAAAGTGAATCTGTCAGACCAATGTGCTAATAGATGTACTAGGAATAAGGACTT 254
Qy 241 ccattcacttgcagggcttttttttttttttttttttttttttttttttttttttttt 300
Db 255 CCATTCTCTCAAGGCTTTGTTTTTTGATAAAGCAAGAAACAAATGCTCTGTTCCCC 314
Qy 301 tcaatagcagtcagagtgaggtagaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 315 TTCAATAGCATGTCAAGTGGAGTGAAGAAATTTGGCCATGAATTTGACCTCTATGAA 374

```

/note="vector: pcwvSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 301 a 152 c 57 g 304 t 30 others
ORIGIN

Query Match 25.0%; Score 432; DB 106; Length 844;
Best Local Similarity 84.5%; Pred. No. 6.5e-106;
Matches 476; Conservative 20; Mismatches 52; Indels 15; Gaps 1;

QY 69 catcgccatccctatgcagaggagacaaagaaagaaatacattcatgaattcaa 128
|||||
DB 1 CATGCCCATCCCTATGCAGAGGGACAAAGGAAAGAAATACAAATTCATGAATTCAA 60
|||||
QY 129 aaatcagcaagacctccctaatacaaatagatccagcactgaagataaaacacaaaa 188
|||||
DB 61 AAATTCAGCAAGACTACCCCTAATCAAAATAGATCCAGCAGTGAAGTAAACCAAAAA 120
|||||
QY 189 agtgaatactgcagaccatgtgctaataagatgtactaggaataaaggacttccattcac 248
|||||
DB 121 AGTGAATACTCAGACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTCCAC 180
|||||
QY 249 ttgcaagcctttgtttgtataagcagaagaaacaaatgcctcgtggtcccttcaatag 308
|||||
DB 181 TTGCAAGCCTTTGTGTTTGTATAAGCAAGAAACAAATGCCTCTGTTGCCCTTCAATAG 240
|||||
QY 309 catgtcaagtggagtgaaaaaagattggcctgaatgttaccctctatgaaaaaacaaga 368
|||||
DB 241 CATGTCAAGTGGAGTAAAAAAGAAATTTKCCATGAATTTTACCTCTATKAANAACAAAGA 300
|||||
QY 369 ctacattagaactgcattgtgtaaggagcgcagctacaaagggaacagtgatctctatcac 428
|||||
DB 301 CTACATTAGAAACTGCATCATTTKGTAAAGGACKSAKSTACATTTTAASATTTATCTATCAS 360
|||||
QY 429 taagagtggcatcaaatgtcagccgtgagttccatgatataccacacgaacac----- 480
|||||
DB 361 TAATATTTTTCATCAAAATTTCTCCCTTTTATTTCCATTTATACCACACTAACACATCTTTT 420
|||||
QY 481 -----agctatcggggtaaagacctacagaaactactgtcgaatcctcgagggga 533
|||||
DB 421 TCSFTCTATCTATCTTTAAATACCTACATAAAACACTACTTCTTAAATCCTCTABTTTA 480
|||||
QY 534 agaaggggaccctggtgtttcacaagcaatccagaggtacgctacgaagtctgtgacat 593
|||||
DB 481 ATAMTTTTTACCCTTTTTCACAAATCAATCCATATTTTACTCTACTAATNCTTTASAT 540
|||||
QY 594 tctcagtggtcagaagttgaat 616
|||||
DB 541 TSSTGATTTTTSATAATTTAAAT 563
|||||

RESULT 7
AL573472/c 899 bp mRNA EST 16-FEB-2001
LOCUS AL573472 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0DI051YC11 3
DEFINITION prime, mRNA sequence.
ACCESSION AL573472
VERSION AL573472.1 GI:12932750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1. (bases 1 to 899)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE JOURNAL COMMENT

Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES source

Location/Qualifiers
1. .899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DI051YC11"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/note="Vector: pcwvSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 199 a 211 c 217 g 248 t 24 others
ORIGIN

Query Match 22.6%; Score 389.2; DB 106; Length 899;
Best Local Similarity 90.1%; Pred. No. 2.4e-94;
Matches 507; Conservative 1; Mismatches 44; Indels 11; Gaps 9;

QY 295 ttcccttcctcaatagcatgtcaagtgaagtgaaagaaatttggccatgaa-tttgacct 353
|||||
DB 736 TGCCCTGTCCCTCATGCTGCATGAGTGAAGAAAGANTTTGGCCATGAANTTTGACCT 677
|||||
QY 354 ctatgaaaaa-aaagactaca-ttagaaactgcattgttgaagaaagcagctacaag 411
|||||
DB 676 CTATGAAGACNAAGACTACANTTAGANACTGTCATCTATTGTT-AAGGACGCGACTACAAG 618
|||||
QY 412 ggaacagtatct--atcactaaagtggtcacaatgctcag--ccctggagttccatgat 467
|||||
DB 617 GGNACAGTATCTATCATCTNAGAGTGGCATCANANTGTCAGCCCTGGAGTNCATGAT 558
|||||
QY 468 accacacgaacacagctatctgggggtaaagacctacagaaactactgtcgaaatcctcg 527
|||||
DB 557 ACCACACGAACACAGCTATTCGGGGTAAAGACCTACAGGAANACNACTGTGAAATCCTCG 498
|||||
QY 528 aggggaagagggggcccttggtttccaagaactccagagggtacgctacgaagtctg 587
|||||
DB 497 AGGGGAAGAGGGGGACCCCTGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTCTG 438
|||||
QY 588 tgacattcctcagtggttcagaagttgaatgcagctgcaatggggagagattatcgagg 647
|||||
DB 437 TGACATTCCTCAGTGTTCAGAAAGTNGAATGCATGACCTGCAATGGGGAGAGTTATCGAGG 378
|||||
QY 648 tctcagtgatcatcacagaatcaggcagagatttgcagcgtggggtatcatcagacaccaca 707
|||||
DB 377 TCTCAT-GATCATACAGAAATCAGGCAAGA-TTGTACGCGCT-GGATCATACAGACACCACA 321
|||||
QY 708 ccggcacaatttcttcctgaaagatatcccgacaagggttttgatgataattattgcgg 767
|||||
DB 320 CCGGCAACAATTTCTGCTGAAAGATATCCCGAAGGGCTTTTGATGATAATTTATGCCG 261
|||||
QY 768 caatcccgatgcccagccagccatgggtgctactactcttgaccctcacaccccgctggga 827
|||||
DB 260 CAATCCGATGGCCAGCCGCGGCTGTTGCTACTTCTTGACCCCTCACACCCCGCTGGGA 201
|||||
QY 828 gtactgtgcaattaaacaatgcy 850
|||||
DB 200 GTACTGTGCAATTAANACATGCG 178
|||||

RESULT 8

AV707491 415 bp mRNA EST 09-OCT-2000
 LOCUS AV707491 ADB Homo sapiens cDNA clone ADBAD07 5', mRNA sequence.
 DEFINITION AV707491
 ACCESSION AV707491
 VERSION AV707491.1 GI:10724756
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 415)
 AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
 TITLE Homo sapiens CDNA ADB clones
 JOURNAL Unpublished (2000)
 COMMENT Contact: Zequang Han
 Chinese National Human Genome Center at Shanghai
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
 201203, P. R. China
 Tel: 86-21-50801919(ex.45)
 Fax: 86-21-50801922
 Email: hanzq@chgc.sh.cn
 This clone is available at CHGC in Shanghai.
 FEATURES
 source
 1. 415
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="ADBAD07"
 /clone_lib="ADB"
 /tissue_type="Adrenal gland"
 /dev_stage="Adult"
 /lab_host="SOLR"
 /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 121 a 106 c 91 g 96 t 1 others
 ORIGIN
 Query Match 9.5%; Score 163.6; DB 32; Length 415;
 Best Local Similarity 94.4%; Pred. No. 1.8e-33;
 Matches 169; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 8 tgacacaaactctgcagccctgtctgtcagcagtctctctgcattctctctctctcc 67
 Db 237 TCACCAAACTCTGCAGCCCTGCTGCTGTGAGCATGTCTCTCTGTCATCTCTCTCTC 296
 QY 68 ccattgcattccctctatgcagaggacaaaagaaagaaagaaatacaattcatgaattca 127
 Db 297 CCATGCCATCCCTATGCAGGGGACAAAGAAAGAAAGAAATACAATTCATGAATTCA 356
 QY 128 aaaaatcagcaagactaccctcaatcaaatagattccagcactgaagataaaacccaaa 186
 Db 357 AAAATCAAGCAAGACTACCTCAATCAAAATAGATCCAGCACTGAAGATTAAACCCAA 415
 RESULT 11
 LOCUS BE244598 EST 15-NOV-2000
 DEFINITION TC8AP2E0846 Pediatric pre-B cell acute lymphoblastic leukemia
 Baylor-HGSC project-TC8A Homo sapiens cDNA clone TC8AP0846, mRNA
 sequence.
 ACCESSION BE244598
 VERSION BE244598.1 GI:9096340
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 321)
 AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project
 Unpublished (2000)
 Contact: Dr. Judith F. Margolin
 Texas Children's Cancer Center and Human Genome Sequencing Center
 at Baylor College of Medicine
 1102 Bates, MC3-3320 Houston, TX 77030, USA
 Tel: 832-824-4536
 Fax: 832-825-4038
 Email: clones@txccc.org
 Seq primer: M13 primer.
 FEATURES
 source
 1. 321
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="TC8AP0846"
 /clone_lib="Pediatric pre-B cell acute lymphoblastic
 leukemia Baylor-HGSC project-TC8A"
 /sex="male"
 /tissue_type="leukopheresis"
 /cell_type="pre-B cell"
 /dev_stage="pediatric 2 years"
 /lab_host="DH10B"
 /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
 First strand cDNA was primed with an anchored
 XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGCGAGGAG(T)VN
 3'; V=A,C,G; N=A,C,G,T] and then dg tailed. Second
 strand was primed with a BamHI-dc primer
 [5'AGAGAGCTCGGATCCGCGCGCAATAATAAT(C) 3'].
 Double-stranded cDNA was then digested with BamHI and
 XhoI and directionally cloned into the BamHI and SalI
 sites of lambda pSB vector. Library went through one
 round of normalization. Library was constructed by Wei
 Yu"
 BASE COUNT 96 a 100 c 59 g 65 t 1 others
 ORIGIN
 Query Match 9.3%; Score 161; DB 165; Length 321;
 Best Local Similarity 99.4%; Pred. No. 8.4e-33;
 Matches 161; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 atgtggtgacaaactctgcagccctgtctgtcagcattctctctctctctctc 60
 Db 160 ATGTGGGTGACCAAACTCCCTGCGAGCCCTGCTGTGTCAGCATGCTCTCTCATCTCTC 219
 QY 61 ctgtctccctctgcccattccctctatgcagaggacaaagaaagaaatacaattcat 120
 Db 220 CTGCTCCCATCGCCATCCCTATGTCAGAGGACAAAGAAAGAAATACAATTCTAT 279
 QY 121 gaattcaaaaatcagcaagactacccttaatacaaatagat 162
 Db 280 GAATTCAAAAATCAGCANAGACTACCCCTAATCAATAATAGAT 321
 RESULT 12
 LOCUS AW175268 mRNA EST 16-NOV-1999
 DEFINITION fi34b04.y1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone
 2639503 5', similar to TR:Q91691 Q91691 GROWTH FACTOR LIVERTINE. ;,
 mRNA sequence.
 ACCESSION AW175268
 VERSION AW175268.1 GI:6441305
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Rasbora; Danio.
 REFERENCE 1 (bases 1 to 661)
 AUTHORS Sugano, S., Kawakami, K., Johnson, S., Li, F., Marra, M., Edgy, S., Hillier, L., Clifton, S., Allen, M., Gibbons, M., Jost, S., Kucaba, T., Martin, J., Pape, D., Steptoe, M., Underwood, K., Theising, B., Ritter, E., Bowers, Y., Wylie, T., Waterston, R. and Wilson, R.

```

VERSION      BF383652.1  GI:11364944
KEYWORDS     EST.
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE     Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL      1 (bases 1 to 766)
COMMENT      NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Jeffrey E. Green, M.D.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM9527 row: o column: 10
              High quality sequence stop: 758.

FEATURES     1. 766
              Location/Qualifiers
              /organism="Mus musculus"
              /strain="FVB/N"
              /db_xref="taxon:10090"
              /clone="IMAGE:4194273"
              /clone_lib="NCI_CGAP_Li9"
              /lab_host="DH10B (T1 phage-resistant)"
              /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
              Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
              Average insert size 1.9 kb. Constructed by Life
              Technologies. Note: this is a NCI_CGAP Library."
              BASE COUNT      171 a 210 c 228 g 157 t
              ORIGIN

Query Match      7.7%; Score 132.8; DB 148; Length 766;
Best Local Similarity 58.3%; Pred. No. 4.6e-25;
Matches 275; Conservative 0; Mismatches 187; Indels 10; Gaps 2;

QY 350 acctctatgaacaaagactacattagaacactgcattgttaaggagcagctaca 409
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 263 ATCTCTCAGAAAGAAATTATGTGCGGACCTGCATATGAGCAATGGGGTCACTACC 322
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 410 agggaaacagtatctactaagaagtgcacaaatgtcagccctgcagttccatgatac 469
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 323 GGGGCACTGTGGCCAGGACAGCTGTGGCCCTGCCCTGCCAAGCCTGGAGTGGAGTTCC 382
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 470 cacacgaacacagctatcggggt-----aaagacctacagaaaactactgtcgaa 520
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 383 CCAATGACCACCAAGTATAGCCGCCACGCCAAAGATGGCTGGAAGAACTCTGTAGGA 442
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 521 atctcgaagggaagaagggggaccctgtgtttcaagaagaatccagaggtacgtacg 580
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 443 ACCCTGATGGGATCCAGAGGTCCTGGTGTCTACTACAAACAAACCGCAGTGTGGTTCC 502
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 581 aagctctgacattcctcagttgttcgaagtgtgaatgcatgacgtgcaatgggagatt 640
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 503 AGAGCTGTGGCATCAAAACCTCGAGGAGGCTGTTGTGTTCTGTGCAACGGTGAGGATT 562
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 641 atcgaagttctatgcatcacagaatcaggcgaagatttgtcagcgtgggatacatcaga 700
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 563 ACCGTGGCAGGTACAGTGTACAGAGTACAGGGCGGGAGTGTACACGTGGGACCTGCAGC 622
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 701 caccacacgggcacaaattcttgctgaagaatatacccgacaggggttttgatgataatt 760
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 623 ACCCCACATCGCACCCCTTTCCAGCCTGAAAGATTCTTAGACAAAGATCTGAAAGACAAC 682
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 761 attgcgcgaatcccgatggccagccgagccatggtgtctatactcttgacc 812
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 683 A-TGTGCTAATCCGACCGAGTCTGAGCGCGCCCTGGTGTCTAAACACACAGACC 733
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 17:44:38 ; Search time 4515.12 Seconds
(without alignments)
5796.398 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atgggggtggtccactoct.....accaccaccaccactag 1692

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:
2: gb_ba2:
3: gb_ba3:
4: gb_in1:
5: gb_in2:
6: gb_in3:
7: gb_om:
8: gb_ov:
9: gb_pat1:
10: gb_pat2:
11: gb_ph:
12: gb_pl1:
13: gb_pl2:
14: gb_pl3:
15: gb_pl4:
16: gb_ba1:
17: gb_ba2:
18: em_fun:
19: em_htgo_hum:
20: em_htgo_inv:
21: em_htgo_rod:
22: em_htg_hum1:
23: em_htg_hum2:
24: em_htg_hum3:
25: em_htg_hum4:
26: em_htg_hum5:
27: em_htg_hum6:
28: em_htg_hum7:
29: em_htg_hum8:
30: em_htg_inv1:
31: em_htg_inv2:
32: em_htg_other:
33: em_htg_rod:
34: em_hum1:
35: em_hum2:
36: em_hum3:
37: em_hum4:
38: em_hum5:
39: em_hum6:
40: em_hum7:
41: em_in:
42: em_om:
43: em_or:

44: em_ov:
45: em_pat:
46: em_ph:
47: em_pl:
48: em_ro:
49: em_sts:
50: em_sy:
51: em_un:
52: em_v1:
53: gb_sts1:
54: gb_sts2:
55: gb_sts3:
56: gb_sy:
57: gb_un:
58: gb_v11:
59: gb_v12:
60: gb_htg1:
61: gb_htg2:
62: gb_htg3:
63: gb_htg4:
64: gb_htg5:
65: gb_htg6:
66: gb_htg7:
67: gb_htg8:
68: gb_htg9:
69: gb_htg10:
70: gb_htg11:
71: gb_htg12:
72: gb_htg13:
73: gb_htg14:
74: gb_htg15:
75: gb_htg16:
76: gb_htg17:
77: gb_htg18:
78: gb_htg19:
79: gb_htg20:
80: gb_htg21:
81: gb_htg22:
82: gb_htg23:
83: gb_htg24:
84: gb_htg25:
85: gb_pr1:
86: gb_pr2:
87: gb_pr3:
88: gb_pr4:
89: gb_pr5:
90: gb_pr6:
91: gb_pr7:
92: gb_pr8:
93: gb_pr9:
94: gb_ro1:
95: gb_ro2:
96: gb_in4:
97: gb_pr10:
98: em_ba3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1692	100.0	1692	9	AX019529 Sequence
2	996.6	58.9	1725	9	AX019527 Sequence
3	829.2	49.0	2219	97	LI1924 Homo sapien
4	829.2	49.0	2232	10	I84668 Sequence 1
5	827.6	48.9	2216	10	E12495 Human cDNA
6	827.6	48.9	2216	10	E12714 Human cDNA
7	827.6	48.9	2216	97	M74178 Human hepat
8	811.6	48.0	2219	10	I36332 Sequence 1

QY 1141 gacctctataaaacaaagactaccattagaaactgcattgtaagaagcagctac 1200
|||||
Db 1141 GACCTCTATGAAACAAAGACTACATTAGAAACTGCATCATTTGTAAGGACGACGATAC 1200
QY 1201 aaggaaacagtattctacataagagtgccatcaaatgtccagccctgagttccatgata 1260
|||||
Db 1201 AAGGAAACAGTATCTATCACTAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCCATGATA 1260
QY 1261 ccacagaacacacagctatcggggtaagagctacaggaataactactgtcgaaatccctcga 1320
|||||
Db 1261 CCACAGAACACACAGCTATCGGGGTAAGAGCTACAGGAAACTACTGTGGAATTCCTCGA 1320
QY 1321 ggggaagaaggagaccctggtgtttcacagcaatccagagggtacgctacgaaagtgtgt 1380
|||||
Db 1321 GGGGAAGAAGGGGACCCTGCTGTTTTCACAAGCAATCCAGAGGTACGCTACGAAAGTCTGT 1380
QY 1381 gacattctcagtggttcagaagttgaatgatcatgacctgcaatggggagagttatcgaggt 1440
Db 1381 GACATTCTCAGTGTTCAGAAGTTGAATTCGATGATGACCTGCAATGGGGAGAGTTATCGAGGT 1440
QY 1441 ctcatgatcatcacagaatcagcgaagattgttcagcgtcggtggtatcatcagacaccacac 1500
|||||
Db 1441 CTCATGGATCATACAGAATCAGCAAGATTGTGAGCGCTGGGATCATCAGACACACAC 1500
QY 1501 cggcacaattcttgctgaaagatatcccgacaaagggttggatgataattatggcgc 1560
|||||
Db 1501 CGGCACAAATTTGCTGTAAGATATCCGACAAAGGCTTTGATGATAATTATTTGGCGC 1560
QY 1561 aatccgatggcagcagcagccatggtgtctatatcttaccctcacaccctgagag 1620
|||||
Db 1561 AATCCCGATGGCAGCGGAGCCATGTTGTATCTCTTGACCCCTCACACCCGCTGGGAG 1620
QY 1621 tactgtgcaattaaacatcgctgacaaagctgacgacgacgacacacacaccacac 1680
Db 1621 TACTGTCCAAATTAACATCGCTGTGACAAAGCTGACGACGACGACAAACACACACCCAC 1680
QY 1681 caccaccactag 1692
Db 1681 CACCACCACCTAG 1692

RESULT 2

AX019527 1725 bp DNA PAT 07-SEP-2000
LOCUS
DEFINITION Sequence 1 from Patent WO9938967.
ACCESSION AX019527
VERSION AX019527.1 GI:10043445
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 1725)
AUTHORS Medico,E., Collesli,C., Comoglio,P., Michieli,P. and Caselli,G.
TITLE Recombinant proteins derived from Hgf and msp
JOURNAL Patent: WO 9938967-A 1 05-AUG-1999;
DOMPE SPA (IT); MEDICO ENZO (IT); COLLESI CHIARA (IT); COMOGLIO
PAOLO (IT); MICHELLO PAOLO (IT); CASELLI GIANFRANCO (IT)
FEATURES
source
1..1725
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="RECOMBINANT"
BASE COUNT 563 a 399 c 390 g 373 t
ORIGIN

Query Match 58.9%; Score 996.6; DB 9; Length 1725;
Best Local Similarity 83.2%; Pred. No. 4.9e-236;
Matches 1166; Conservative 0; Mismatches 214; Indels 21; Gaps 2;

QY 292 tgtgacctcttcagaagaagactacgtacgtgagcctgcattcgaacaatgggggtggg 351
|||||

Db 346 TTTGACCTCTATGAAACAAAGACTACATTAGAAACTGCATCATTTGGTAAAGGACGACG 405
QY 352 taccggggacaccatggccacgaccgtgggtggcctgcctgcacgagcttggagccacaag 411
|||||
Db 406 TACAAGGGAACAGTATCTATCACTAAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCCATG 465
QY 412 ttccggaatgatcacaaagtacacgcccactctccggaatggcctggaagaactctgc 471
|||||
Db 466 ATACCACACGAACACAGCTA-----TCGGGGTTAAAGACTACAGGAAAACTACTGT 516
QY 472 cgtaacctctgatggcgaccccgaggtccttggtctacacaaacagacctgctgctgcgc 531
|||||
Db 517 CCAAAATCTCGAGGGGAGAGAGGGGACCCTGCTGTTTTCACAAGCAATCCAGAGGTACGC 576
QY 532 ttccgagagctcgcgcatcaaatccctgcggggagggcgctgtgctgggtgcaatggcgag 591
|||||
Db 577 TACGAAGTCTGTGACATTCCTCAGTGTTCAGAAAGTTGAATGCATGACCTGCAATGGGGAG 636
QY 592 gaatacccgggcgcggttagaccgcgcgagtcaggggcgagtgccagcgctggatctt 651
|||||
Db 637 AGTTATCGAGGTCTCATGGATCATACAGAAATCAGGCAAGATTGTGAGCGCTGGGATCAT 696
QY 652 cagcaccgcgacccacaccccttcgagccgcgcaagtctctcgaccaaaggtctggacgac 711
|||||
Db 697 CAGACACACACCCGCGACAAATTTCTGCTGAAGATATCCCGACAGGGCTTTGATGAT 756
QY 712 aactattccggaatctctgacggctccgagcggtccatggtctacactacggtccgcgag 771
|||||
Db 757 AATTTATGCGGCAATCCGATGGCCAGCGGAGGCCATGCTGCTATACTCTTGACCCCTCAC 816
QY 772 atcgagcgagagttctgtgacctcccccgctcggggtcgagggcacagcccccgctcgag 831
|||||
Db 817 ACCCCCTGGGAGTACTGTGCAATTAACCATCGCTGACAAAGC-----TTGC 864
QY 832 ggcggtggcggtctctggtggcggtggctcgcggtggcggttctctagagggacaaag 891
Db 865 GCGGTGGCGGTTCTGTTGGCGGTGGCTCCGCGGTGGCGGTTCTCTAGAGGGACAAAGG 924
QY 892 aagaagaataacattcatgaattcaaaaaatcagcaaaagactacccttaataaata 951
Db 925 AAAAGAAGAAATACAAATTCATGAATTCAAAAATCAGCAAGACTACCCCTAATCAAAATA 984
QY 952 gatccagcactgaagataaaaaacccaaagtgaatactactgcagacccaatgtgctaata 1011
Db 985 GATCCAGCAGCTGAAGATAAAAAACCAAAAGTGAATTAAGTGCAGACCAATGTCTAATA 1044
QY 1012 tctactagaataaaagagcttccattcacttcaaggttttcttcttataaagaaga 1071
Db 1045 TGTACTAGAAATAAAGGACTTCCATTCACTTGCAGGCTTTGTTTGTGATAAAGCAGA 1104
QY 1072 aaacaatgctcgtggtcccttcaatagctatgctcaagtggagtgaaaaaagatttggc 1131
Db 1105 AAACAATGCTCTGTTCCCTTCAATAGCATGTCAAGTGGAGTGAATAAAGAAATTTGGC 1164
QY 1132 catgaattgacctctatgaaaaacaaagactacattagaaactgcatcattgttaaagga 1191
Db 1165 CATGAATTTGACCTCTATGAAAAACAAAGACTACATTAGAAACTGCATCATTTGTTAAAGGA 1224
QY 1192 cgcagctcaaggaacagttatctatcaataagagtggtcaaatgtcagccctggagt 1251
Db 1225 CCGAGCTACAGGGGAACAGTATCTATCACTAGAGTGGCATCAATGTTCAGCCCTGGAGT 1284
QY 1252 tccatgataccacacgaacacagctatcggggttaagacctacagaaaaactactgtcga 1311
Db 1285 TCCATGATACCACACGACACAGCTATCGGGTTAAGACCTCAGAGAAACTACTGTGCA 1344
QY 1312 aatctcaggggaagaagggggacctggtgtttcacaaagcaatccagagggtacgtac 1371
Db 1345 AATCTCGAGGGGAAGAAGGGGACCCTGCTGTTTTCACAAGCAATCCAGAGGTACGCTAC 1404
QY 1372 gaagctgtgacattctcagttcagagttgaatgcatgacctgcaatggggaagt 1431
Db 1405 GAAGTCTGTGACATTCCTCAGTGTTCAGAAAGTTGAATGATGACCTGCAATGGGGAGAGT 1464

184668
LOCUS 184668 2232 bp DNA PAT 04-APR-1998
DEFINITION Sequence 1 from patent US 5696086.
ACCESSION 184668
VERSION 184668.1 GI:3022188
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2232)
AUTHORS Avraham,H,Karsenty and Godowski,P.J.
TITLE Methods and kits using macrophage stimulating protein
JOURNAL Patent: US 5696086-A 1 09-DEC-1997;
FEATURES
Location/Qualifiers
1..2232
/organism="unknown"
BASE COUNT 448 a 672 c 686 g 426 t
ORIGIN

Query Match 49.0%; Score 829.2; DB 10; Length 2232;
Best Local Similarity 99.6%; Pred. No. 1.4e-194;
Matches 831; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 atgggtggtcccaactcctgctctgactcctgaactaacttaggggtccctggcagcgc 60
Db 11 ATGGGGTGGCTCCCACTCCTGCTGCTTCTGACTCAATGCTTAGGGGTCCCTGGGAGCGC 70

Qy 61 tcgccattgaatgacttccaagtgtccggggcagagctacagcactgctacatgcg 120
Db 71 TCGCCATTGAATGACTTCCAGTGCTCCGGGCGCAGAGCTACAGCACCCTGTATACGG 130

Qy 121 gtggtgcccgggcttggcaggagatgtgcagatgctgaagatgtgctggtgcgtgt 180
Db 131 GTGTGTCCCGGCGCTTGGCAGAGGATGTGCAGATGCTGAAGAGTGTCTGTGCTGT 190

Qy 181 gggcccttaattgactgcgggcttcactacacagtgagcagcatgttgcacactg 240
Db 191 GGGCCCTTAATGGAGTGGCGGCGCTTCCACTACACGCTGAGCAGCATGTTGCCAATG 250

Qy 241 ctgccatgactcaactcgcggccacacagagctgcggcgtctctggcgctgtgacctc 300
Db 251 CTGCCATGGACTCACATCGCCCGCCACAGAGGCTGCGCGCTTCTGGGCGCTGTACCTC 310

Qy 301 ttccagaagaagactacgtacggacctgcatcgaacaatgggtgggtaccggggc 360
Db 311 TTCCAGAAGAAGACTACGTACGGACCTGCATCATGAACAATGGGTTGGTACCGGGC 370

Qy 361 accatggccagcagcgtgggtggcctgcctgcagagcttgagccacaaagtcccgaa 420
Db 371 ACCATGGCCAGCAGCGTGGGTGGCGTGCCTGCCAGGCTTGGAGCCACAAAGTCCCGAAT 430

Qy 421 gatacaagcacacccacactccggaatggcctgggaagagaactctgcctgaacct 480
Db 431 GATCACAGTACACGCCACACTCTCGGGAATGGCTTGGAGAGAATCTTGCCTGTACCT 490

Qy 481 gatggcagcccgagggtccttgggtgtacacaaacagaccctgctgctgctccagagc 540
Db 491 GATGCGGACCCCGGAGGTCTTGGTGCTACACACAGACCCTGCTGTGGCTTCCAGAGC 550

Qy 541 tcgggcatcaaatcctgcggaggccgcgtgtgctcctgggtgcaatggcgagaaataccgc 600
Db 551 TCGCGCATCAAAATCTGCGGGAGGCGCGTGTGTCTGTGTAATGGCGAGGAATACCGC 610

Qy 601 ggcggttagaccgacagagtagcggcgagtgctccagcgtggtatcttcagaccgc 660
Db 611 GCGCGGTAGACCGCACGGAGTACAGGGCGAGTGCCAGCGCTGGGATCTTCAGCACCG 670

Qy 661 caaccagcccttcgagcgggcaagtctctcgaccaggtcctggagacaactattgc 720
Db 671 CACAGACCCCTTCGACCGCGGCAAGTCTTCGACCAGGTTCTGGACGACAACTATTGC 730

Qy 721 cggaatctgacgggtccgagcggccatgtgtgtacactacggtatccgcagatcgagcga 780

Db 731 CGGAATCTTACGGCTCCGAGCGCCATGGTGTACACTACGATCCGACAGATCGACGA 790
Qy 781 gagtctgtgacctccccgcgtgcgggtccgaggccagccccgcctcgagggc 834
Db 791 GAGTCTGTGACTCCCGCTCGGGTCCGAGGCACAGCCCGCAAGAGGCC 844

RESULT 5
E12495 2216 bp DNA PAT 24-JUN-1998
LOCUS Human cDNA encoding macrophage stimulating protein.
DEFINITION
ACCESSION E12495
VERSION E12495.1 GI:3251328
KEYWORDS JP 1997003099-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2216)
AUTHORS Yoshikawa,W., Shimonishi,M., Iwamoto,J., Takehara,T. and Hagiya,M.
TITLE MODIFIED MATERIAL OF MACROPHAGE STIMULATING PROTEIN AND ITS
JOURNAL
COMMENT
TOYODO CO LTD
OS Homo sapiens (human)
PN JP 1997003099-A/1
PD 07-JAN-1997
PF 20-JUN-1995 JP 1995153309
PI YOSHIKAWA WATARU, SHIMONISHI MANABU, IWAMOTO JUNKO, PI
TAKEHARA TOYOHIRO,
PI HAGIYA MICHIO
PC C07K14/535,C07H21/04,C12N5/10,C12N15/09,C12P21/02,(C12N5/10,
C12R1:91),
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..2216
/organism="Homo sapiens"
/cell_line="Hep G2"
FT CDS 1..2136
/product="macrophage stimulating protein" FT
FT sig_peptide 1..54.
Location/Qualifiers
source 1..2216
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 445 a 566 c 682 g 423 t
ORIGIN

Query Match 48.9%; Score 827.6; DB 10; Length 2216;
Best Local Similarity 99.5%; Pred. No. 3.4e-194;
Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 atgggtggtcctccactcctgctgcttctgactcaatgcttaggggtccctggcagcgc 60
Db 1 ATGGGGTGGCTCCCACTCCTGCTGCTTCTGACTCAATGCTTAGGGGTCCCTGGGAGCGC 60

Qy 61 tcgccattgaatgacttccaagtgtccggggcagagctacagcactgctacatgcg 120
Db 61 TCGCCATTGAATGACTTCCAAAGTGTCCGGGCGCAGAGCTACAGCACCCTGTATACGG 120

Qy 121 gtggtgcccgggcttggcaggagatgtggcagatgctgaagatgtgctggtcgctgt 180
Db 121 GTGTGTCCCGGCGCTTGGCAGAGGATGTGCAGATGCTGAAGAGTGTGTGTGCTGT 180

Qy 181 gggcccttaattgactgcgggcttcactacaaagtgcagcaccatgggtgcgaactg 240
Db 181 GGGCCCTTAATGGAGTGGCGGCGCTTCCACTACAACTGAGCAGCCATGTTGCCAATG 240

HUMHEPFA HUMHEPFA 2216 bp mRNA PRI 07-OCT-1994
 LOCUS Human hepatocyte growth factor-like protein mRNA, complete cds.
 DEFINITION M74178
 ACCESSION M74178.1 GI:183976
 VERSION hepatocyte growth factor-like protein; serine protease-like domain.
 KEYWORDS Homo sapiens cDNA to mRNA.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 2216)
 AUTHORS Han, S., Stuart, L.A., and Degen, S.J.
 TITLE Characterization of the DNL1582 locus on human chromosome 3: identification of a gene coding for four kringle domains with homology to hepatocyte growth factor
 JOURNAL Biochemistry 30 (40), 9768-9780 (1991)
 MEDLINE 92002016
 REFERENCE 2 (bases 1 to 2216)
 AUTHORS Degen, S.J.
 TITLE Direct Submission
 JOURNAL Submitted (02-DEC-1991) S.J.F. Degen, Division of Basic Science Research, Children's Hospital Research Foundation, Cincinnati, OH 45229-3039, USA
 FEATURES
 Location/Qualifiers
 1..2216
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stage="fetal"
 /chromosome="3"
 /map="3p21"
 /tissue_type="liver"
 1..2136
 /codon_start=1
 /product="hepatocyte growth factor-like protein"
 /protein_id="AA50165.1"
 /db_xref="GI:183976"
 /translation="MGWPLLLLLTOYLGVPQSRPLNDFQVLRGTEQLHLLHAVPG PWQEDVAEEACAGCGPLMDCRAPHYVNSHGGCOLLPWTOHSPHRLRRSRQDLFQ KDQYRTIMNNGVYRGTMATTVGGLPQAWSHKFPNDHKYTPTRNGLEENFCRNP DGDGPGWCYTTFDAPQSGQIKSCREAACVWCNGEYRGAVDRTEGRCQWDLO HPQHPFPFGFLDQGLDNDYCRNPDGSRWPWCYTDPQIEREFCDLPRGSEAOPRQ EATVSCFRGKGEYRGANTTTAGVPCQRDAQIAPHQRTTPEYACKDLRENFCRN PDGSAEPFTLRPMRAFCYIRRCTDDVRPQCYHGAGEQYRGTVSKTRKGVOCC RWSATPHKPOFTFSEHAOLEENFCRNPDSHGSPWCYTMDPTDFDYCALRRCAD DQPSILDPQVQPEKCGKVRDLDRKSLRVVGHGPNPWTSLNRQGHFCG GSLVKEOIMLTAROCFSCHMLPTGYEVLGLTFONPOHGEPSLQRPVAKWVCGPSG SOLVLLKLSRSTLNRVALICLPEWVVPPTKCEIAGWETKGTGNDTVLNVALL NVISNQCENIKHRGRVRESEMTEGLLAPVACGEDYGGPLACFTHNCWVLEGILIPN RVCARSRPWFTRVSVFVDMHVKVRLG"
 328..558
 /note="first kringle domain"
 571..804
 /note="second kringle domain"
 847..1083
 /note="third kringle domain"
 1108..1344
 /note="fourth kringle domain"
 1449..1450
 /note="activation site"
 1450..2133
 /note="serine protease-like domain"
 2187..2192
 polyA_signal
 BASE COUNT 445 a 666 c 682 g 423 t
 ORIGIN

Query Match 48.9%; Score 827.6; DB 97; Length 2216;
 Best Local Similarity 99.5%; Pred. No. 3.4e-194;
 Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 1 atgggggtgctccactctgtctgtactcaatgtttgaggggtccctggcgagcgc 60
 Db 1 ATGGGGTGGCTCCCACTCTGTCTGTCTGACTCAATGTCTGGGGTCCCTGGGCGCGC 60

Qy 61 tggccattgaatgaacttccaagtgtcccgggggcacagagctacagacacctgtacatgcg 120
 Db 61 TGGCCATTGAATGAACTTCCAAAGTGTCCGGGGCACAGAGCTACAGACCTGCTACATGCG 120
 Qy 121 gtagtcccgggccttgaggagagatgtagcagatgctgaagatgctggtcgtgt 180
 Db 121 GTAGTCCCGGGCCTTGAGGAGGATGTGCAGATGCTGAAGATGCTGTGTGCTGCTGT 180
 Qy 181 gggcccttaagtgaactgcccggggtctccactacaacgtagcagcagccatggtgccaaactg 240
 Db 181 GGGCCCTTAATGAGACTGCGGGGCTTCCACTACAACGTGAGCAGCCATGTGTGCCAACTG 240
 Qy 241 ctgcaatgactcaaacactgccccacacagagctggcggttcttggcgctgtgacctc 300
 Db 241 CTGCCATGGACTCAACACTCGCCCCACACGAGGCTGCGGCGCTTCTGGGCGCTGTGACCTC 300
 Qy 301 ttccagaagaagactacgtacggaacctcatcatgaacaatggggttgggtaccggggc 360
 Db 301 TTCCAGAAGAAGACTACGTACGAGACTGCTCATGATGACCAATGGGTGGGTACCGGGGC 360
 Qy 361 accatggcacgacctgggtggcctgccagcttggagccttggagcacaagttcccgaaat 420
 Db 361 ACCATGGCACGACCGTGGGTGGCTGCCCTGGCAGGCTTGGAGCCACAAGTTCCCGAAT 420
 Qy 421 gatcacaagtacacgcccactctccggaatggcctgggaagagaacttctccgtaacct 480
 Db 421 GATCACAAGTACACGCCCACTCTCCGGAATGGCTTGAAGAGAACTTCTCCCGTAACCT 480
 Qy 481 gatggcagccccgaggtccttgggtgtgtacacacagacctgtgtgcttccagagc 540
 Db 481 GATGGCAGCCCCGAGGTCTTGTGTGTACAAACAGACCTGCTGTGCGCTTCCAGAGC 540
 Qy 541 tgcggcatcaaatcctgcgggggagcgctgtgtctgttgcgaatggcgaggaataccgc 600
 Db 541 TCGCGCATCAATCCTGCGGGAGGCGCGCTGTCTGTGTCATGTCGAGGAATACCGC 600
 Qy 601 ggcgcgttagaccgcagcgagtcaggcgcgagtcgcagcgctgggagtcttcagacccgc 660
 Db 601 GCGCGGTAGACCGCACCGAGTCAGGCGCGAGTGCAGCGCTGGGATCTTTCAGACCCG 660
 Qy 661 caccagcacccttcagcggcggaagtctcagcaaatctcagcaaatgctggacgacaattgc 720
 Db 661 CACCAGCACCCCTTCGAGCGCGGCAAGTCTCTGACCAAGGCTTGGACGACAACTATTCG 720
 Qy 721 cggaaatcctcagcgtcgcagcgccatgtgtctacactacgcatccgagatccgagcga 780
 Db 721 CGGAATCTTGACGCTCCGAGCGGCCATGTGTGTACTACTAGGATCCGCAGATCGCGGA 780
 Qy 781 gaggcttgacctccccctgcgggtccgagggcacagcccccgctcgagggc 834
 Db 781 GAGTCTGTGACCTCCCGCGCTGCGGGTCCGAGGCACAGCCCGCCCAAGAGGCC 834

RESULT 8
 LOCUS I36332 2219 bp DNA PAT 13-MAY-1997
 DEFINITION Sequence 1 from patent US 5606029.
 ACCESSION I36332
 VERSION I36332.1 GI:2086845
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 2219)
 AUTHORS Degen, S.J.
 TITLE Gene for a growth factor and its cDNA and protein
 JOURNAL Patent: US 5606029-A 1 25-FEB-1997;
 FEATURES Location/Qualifiers
 1..2219
 /organism="unknown"
 BASE COUNT 461 a 660 c 675 g 419 t 4 others
 ORIGIN

I36333
LOCUS I36333 2219 bp DNA PAT 13-MAY-1997
DEFINITION Sequence 2 from patent US 5606029.
ACCESSION I36333
VERSION I36333.1 GI:2086846
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2219)
AUTHORS Degen,S.J.
TITLE Gene for a growth factor and its cDNA and protein
JOURNAL Patent: US 5606029-A 2 25-FEB-1997;
FEATURES
Location/Qualifiers
1..2219
/organism="unknown"
BASE COUNT 461 a 660 c 674 g 420 t 4 others
ORIGIN

Query Match 47.9%; Score 810; DB 10; Length 2219;
Best Local Similarity 99.4%; Pred. No. 7.7e-190;
Matches 813; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 17 tccctgtctctgactcaatgcttaggggtccctggcagcgctgcgcattgaatgact 76
Db 1 TCCTGCTCTTCTGACTCAATACTTAGGGGTCCCTGGGCAGCGCTCGCATTTGAATGACT 60

Qy 77 tccaaagtctcggggcacagactacagcacctgctacatgcggtgctccggcctt 136
Db 61 TCCAAAGTCTCCGGGGCACAGACTACAGCACCTGCTACATGCGGTGTGTCGGGCGCTT 120

Qy 137 ggcaggagatgtgacagatgctgaagagtgtgctggtgcgtgtggtgccccttaattggact 196
Db 121 GGCAGGAGATGTGCACATGCTGAAGAGTGTGCTGCTGCTGTGGGCCCTTAATGGACT 180

Qy 197 gccgggcccctccatacaaacgtgagcagccatggttgccaaactgtgcattggaactcaac 256
Db 181 GCCGGGCTTCCACTACAACTGAGCAGCCATGTTGCCAACTGCTGCCATGGACTCAAC 240

Qy 257 actgcggccacacagagctgcggcttctggcgactgacactcttcagagaagaagact 316
Db 241 ACTCGCCCCACAGAGGCTCGGGCTTCTGGCGCTGTGACTCTTCCAGAGAAAGACT 300

Qy 317 acgtacgagacctgcatcatgaacatgggttgggtggtacccggggcaccatggccacgacgg 376
Db 301 ACGTAGCGACCTGCATCATGAACATGGGTTGGGTACCGGGGCACCATGGCCACGACGG 360

Qy 377 tgggtggcctgcctgcaggtctggagccacaagtccccaagtccccaagtatcacagtacacgc 436
Db 361 TGGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAGTTCGCCGAATGATCACAAGTACAGGC 420

Qy 437 ccactctccgaatggcctgaaagaactctccgttaaccctgactgagcaccgccgag 496
Db 421 CCACTCTCCGGAATGGCTGGAAGAAGAACTTCTGCCGTAAACCCCTGATGGCAGCCCGGAG 480

Qy 497 gtccttggtgtacacaacagacctgtgtgcgtcttcagagctggcgcatcaaatcct 556
Db 481 GTCCTTGGTGTACACACACACACCTGCTGTGCGCTTCCAGAGTGGCGATCAATCCT 540

Qy 557 gccggggagcccgctgtgtgtgtgcaatggtgcgaggaataccggggcggttagaccgca 616
Db 541 GCCGGGAGCGCGCTGTGCTGTGTGCAATGGCAGGAATACCGGGCGCGGTAGACCCGA 600

Qy 617 cgaagtcaaggcgagtgccagcgtgggattcttcagcaccgccaccagcacccttcg 676
Db 601 CGGAGTCAAGGCGGAGTTCCAGCGCTGGGATCTTCAGACCCCGCACAGACCCCTTCG 660

Qy 677 agccgggcaagtctctgaccacaaggtctggcgacacaactatttcgggaattcctgacgct 736
Db 661 AGCGGGCAAGTCTCTGCACCAAGGCTGGACGACAACTATTGCGGGAATCCTGACGCT 720

Qy 737 ccgagcgccatggtgtactactacggaatccgcagatcgagcagaggttctgtgacctcc 796

Db 721 CCGAGGGCCATGGTGTCTACTACATCGGATCCGAGATCGAGCGAGTCTGTGACTCC 780
Qy 797 cccgctgcgggtccgagggcacagcccccgcctcgagggc 834
Db 781 CCCGCTGCGGGTCCGAGGCACAGCCCCGCCAAGAGGCC 818

RESULT 11
LOCUS I36334 2021 bp DNA PAT 13-MAY-1997
DEFINITION Sequence 3 from patent US 5606029.
ACCESSION I36334
VERSION I36334.1 GI:2086847
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2021)
AUTHORS Degen,S.J.
TITLE Gene for a growth factor and its cDNA and protein
JOURNAL Patent: US 5606029-A 3 25-FEB-1997;
FEATURES
Location/Qualifiers
1..2021
/organism="unknown"
BASE COUNT 426 a 609 c 612 g 374 t
ORIGIN

Query Match 46.9%; Score 794.2; DB 10; Length 2021;
Best Local Similarity 99.6%; Pred. No. 6.3e-186;
Matches 796; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 36 atgcttaggggtccctggcagcgctgcgcattgaatgacttccaaagtgtccggggcac 95
Db 1 ATGCTTAGGGGTCCCTGGCAGCGCTCGCCATTGATGACTTCCAAAGTGTCTCCGGGCAC 60

Qy 96 agagctacagcacctgtctacatgcgtggtgcccgggcttggcaggagatgtggcaga 155
Db 61 AGAGCTACAGCACCTGCTACATGCGGTGTGCGCGGCCCTTGGCAGGAGATGTGGCAGA 120

Qy 156 tgctgaagagtgtgctgcgtgtgggccccttaaggactgcgggccccttcactacaa 215
Db 121 TCCTGAAGAGTGTGCTGCTGTGGGCCCTTAATGGACTGCCGGCCCTTCCACTACAA 180

Qy 216 cgtgagcagccatggttgcacactgctgcatgactcaacactcgcccacacagagct 275
Db 181 CGTGAGCACCATGTTGTCACACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

Qy 276 gcggcgttctggcgctgtgacctcttcagagaagaagactacgtacggacctgcacat 335
Db 241 CGGGCGTCTGGCGCTGTGACCTCTTCCAGAGAAAGACTACGTACGGACCTGCATCAT 300

Qy 336 gaacaatggggttgggtacccggggcaccatggccagaccgtgggtggcctgcctgcca 395
Db 301 GAACAATGGGTTGGGTACCGGGGCACCATGGCAGACCGTGGGTGGCCTGCCCTGCCA 360

Qy 396 ggttggagccacaagtcccaaatgatacacaagtacagccacactctccggaatggcct 455
Db 361 GCCTTGGAGCCACAAAGTTCGGAATGATCACAAGTACACGCCACACTCTCCGGGAATGGCCT 420

Qy 456 ggaagagaacttctgcgttaaccctgattggcaccgccggaggctccttgggtgctacacaac 515
Db 421 GGAAGAGAACTTTCGCCGTAAACCTGATGGCAGACCCCGGAGGTCTTGGTGTACACAAC 480

Qy 516 agaccctgctgtgccttccagagtgcggcatcaaatcctgcggagagccgctgtgt 575
Db 481 AGACCCTGTGCTGCGCTTCCAGAGCTGCGGCATCAAAATCTCGCGGAGGCCCGCTGTGT 540

Qy 576 ctgggtgcaatggcaggaataccggcgcggttagaccacacaggaatcagggcgcagtg 635
Db 541 CTGGTGCAAATGGCAGGAATACCGCGGCGGTAGACCCGCGGAGTACAGGGCGCGAGTG 600

Db	262	GTTTTGTAAACGACGAGAAACAATGCCTCTGGTTCCTCCCTTCAATAGCATGTCAAGTGA	321
Qy	1114	gtgaaaaagaatttgccatgaatttgacctctatgaacacaaagactacattagaac	1173
Db	322	GTGAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAACAAACAGACTTACATTAGAAC	381
Qy	1174	tgcatttggttaaaaggacgcagctacaagggaacagtagtctcactaagaagtggcctc	1233
Db	382	TGCATCATTTGTTAAAGGACGACAGCTACAAGGGAACAGTATCTATCACTAAGAGTGGCATC	441
Qy	1234	aaatgtcagccctggaggttcctatgataccacacgacacagctatcg99gttaaacctta	1293
Db	442	AAATGTGAGCCCTGGAGTTTCATGATATACACACACGACACAGCTATCGGGGTAAAGACCTA	501
Qy	1294	caggaaaactactgtcgaaatcctcgaagggaaggaaggagccctggtgtttcacaagc	1353
Db	502	CAGGAAAACACTACTGTGCAAAATCCTCGAGGGGAAGAGGGGGACCTGGTGTTCACAAAGC	561
Qy	1354	aatccagaggtacgctacgaagtctgtgacattcctcagttctcagaagtgtgaatgcctg	1413
Db	562	AATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAGTGTTCAGAAGTTGAATGCATG	621
Qy	1414	acctgcaatgggagaggttatcgaggtctcattggtatcacagaatcagggcaagattcgt	1473
Db	622	ACCTGCAATGGGAGAGTTTATCGAGGTCCTATGATATACAGAAATCAGGCAAGATTTGT	681
Qy	1474	cagcgtcgagatcatcagacacacacccgacaaattcttcgctgaagatatcccgac	1533
Db	682	CAGCGCTGGGATCATCAGACACACACCCGGCACAAATCTTCGCTGAAAGATATCCCCGAC	741
Qy	1534	aaggccttgatgaataattatggcccaatcccgatggccagccgagggccatgggtgctat	1593
Db	742	AAGGCTTTGATGATAATTATTGGCCGAATCCGATGCGCAGCCGAGGCCATGGTGCTAT	801
Qy	1594	actctgacctcacaccgcctgggagtagtctgcaattaaacatcgctgcacaaagt	1653
Db	802	ACTCTTTGACCTCACACCCGCTGGGAGTACTGTGCAATTTAAACATGCGCTGACAATACT	861
Qy	1654	gacgacac 1662	
Db	862	ATGAATGAC 870	
RESULT	14		
LOCUS	E05008		
DEFINITION	HLC2 cDNA sequence coding for human hepatoparenchymal cell growth factor(HGF).		
ACCESSION	E05008		
VERSION	E05008.1	GI:2173203	
KEYWORDS	JP 1993111383-A/2.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Nakamura,T., Hagiya,M., Seki,T., Shimonishi,M., Shimizu,S., Inohara,I., Sakaguchi,M. and Asami,O.		
TITLE	RECOMBINANT HUMAN HEPATOCYTE GROWTH FACTOR AND ITS PRODUCTION		
JOURNAL	Patent: JP 1993111383-A 2 07-MAY-1993;		
COMMENT	NAKAMURA TOSHIIICHI OS Homo sapiens (human) PN JP 1993111383-A/2 PD 07-MAY-1993 PF 06-JUN-1991 JP 1991163485 PR 11-JUN-1990 JP 90P 152474 PI NAKAMURA TOSHIIICHI, HAGIYA MICHIO, SEKI TATSUYA, PI SHIMONISHI MANABU, PI SHIMIZU SHIN, INOHARA IZUMI, SAKAGUCHI MARIKO, ASAMI OSAMU C12N15/16,C07K13/00,C12N5/10,C12N15/85,C12P21/02//A61K37/02, PC A61K37/24, CC (C12P21/02,C12R1:91); strandedness: Double;		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 18:33:25 ; Search time 214.84 Seconds
(without alignments)
4945.120 Million cell updates/sec

Title: US-09-600-991-3

Perfect score: 1692

Sequence: 1 atggggtggtccactctct.....accaccaccaccactag 1692

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

```
1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1692	100.0	1692	AA087663	HGF-MSP hybrid pro
2	996.6	58.9	1725	AA087662	Hepatocyte growth
3	827.6	48.9	2213	AA047149	Macrophage stimula
4	827.6	48.9	2216	AA047145	Macrophage stimula
5	827.6	48.9	2216	AA072083	Human MSP cDNA. H
6	811.6	48.0	2219	AA079723	Human L5/3 tumour
7	811.6	48.0	2219	AA079723	Human L5/3 partial
8	811.6	48.0	2262	AA079729	Full-length human
9	811.6	48.0	2262	AA079729	Human L5/3 partial
10	810	47.9	2219	AA079724	Human L5/3 tumour
11	810	47.9	2219	AA079724	Human L5/3 partial

12	794.2	46.9	2021	18	AA062438	Human L5/3 partial
13	792.6	46.8	2021	15	AA079725	Human L5/3 tumour
14	773	45.7	2190	20	AA087676	HGF-MSP hybrid pro
15	773	45.7	2190	20	AA087677	HGF-MSP hybrid pro
16	773	45.7	2199	12	AA015177	Human leukocyte-de
17	773	45.7	2199	11	AAV53627	Human leukocyte-de
18	773	45.7	2289	11	AAQ06088	Tumour cytotoxic f
19	771.4	45.6	2172	19	AAV61952	Nucleotide sequenc
20	771.4	45.6	2173	14	AAQ37308	Encodes haematopoi
21	771.4	45.6	2289	13	AAQ21066	TCF II-encoding se
22	766.8	45.3	2172	12	AAQ14182	Plasminogen-like g
23	748	44.2	2184	13	AAQ20049	Human hepatocyte g
24	748	44.2	2184	14	AAQ46040	Hepatocyte growth
25	748	44.2	2187	14	AAQ47832	Competative inhibi
26	748	44.2	2187	14	AAQ47833	Competative inhibi
27	748	44.2	2187	14	AAQ45702	Vascular endotheli
28	748	44.2	2214	12	AAQ15176	Human leukocyte-de
29	748	44.2	2214	19	AAV53626	Human leukocyte-de
30	746.4	44.1	2184	15	AAQ56157	Hepatocyte growth
31	746.4	44.1	2187	12	AAQ10489	Hepatic parenchyma
32	746.4	44.1	2393	12	AAQ14038	Human hepatocyte g
33	743.2	43.9	2184	13	AAQ26052	Human HGF. Homo s
34	739.2	43.7	873	13	AAQ23662	Truncated hepatocy
35	739.2	43.7	873	14	AAQ34613	cDNA encoding 34kD
36	739.2	43.7	873	18	AA078409	34 kilodalton hepa
37	735.8	43.5	1661	12	AAQ12398	Human hepatocyte g
38	735.8	43.5	1661	13	AAQ22144	Human HGF gene par
39	735.8	43.5	2187	12	AAQ12399	Human hepatocyte g
40	735.8	43.5	2187	13	AAQ22146	Complete human HGF
41	735.8	43.5	2187	19	AAV20546	Human recombinant
42	735.8	43.5	2187	20	AA015186	Nucleic acid encod
43	615.8	36.4	1400	13	AAQ20745	Encodes alpha-chai
44	614.2	36.3	2187	13	AAQ20255	Encodes alpha- and
45	564.2	33.3	2216	20	AAV72084	Mouse MSP cDNA. M

ALIGNMENTS

RESULT	1				
AA087663					
ID	AA087663	standard; cDNA; 1692 BP.			
XX	AA087663;				
AC	AA087663;				
XX	26-OCT-1999	(first entry)			
DT	26-OCT-1999	(first entry)			
XX	HGF-MSP hybrid protein Metron Factor-1 cDNA.				
DE	Metron Factor-1; human; hepatocyte growth factor; HGF;				
XX	macrophage stimulating proteinl MSP; apoptosis; chemotherapy;				
KW	toxicity; therapy; ss.				
XX	Homo sapiens.				
OS	Synthetic.				
XX	Key	Location/Qualifiers			
PH	sig_peptide	1..54			
FT	mat_peptide	/*tag= a			
FT	mat_peptide	55..1692			
FT	mat_peptide	/*tag= b			
XX	W09938967-A2.				
PN	W09938967-A2.				
XX	05-AUG-1999.				
PD	05-AUG-1999.				
XX	27-JAN-1999;	99WO-EP00478.			
PF	27-JAN-1999;	99WO-EP00478.			
XX	30-JAN-1998;	98IT-MI00179.			
PR	30-JAN-1998;	98IT-MI00179.			
XX	(DMP-) DOMPE SPA.				
PA	(DMP-) DOMPE SPA.				
XX	Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;				
PI	Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;				

Qy 1561 aatccgatgccagccagccatggtgctatactcttgacctcacaccgctggag 1620
 Db |||||||
 Qy 1561 aatccgatgccagccagccatggtgctatactcttgacctcacaccgctggag 1620
 Db |||||||
 Qy 1621 tactgtgaattaaacatgcgtgacaaagctgacgacgacgacaaacacaccacac 1680
 Db |||||||
 Qy 1621 tactgtgaattaaacatgcgtgacaaagctgacgacgacgacaaacacaccacac 1680
 Db |||||||
 Qy 1681 caccaccactag 1692
 Db |||||||
 Qy 1681 caccaccactag 1692
 Db |||||||

RESULT 2

AAX87662
 ID AAX87662 standard; cDNA; 1725 BP.

AC AAX87662;
 XX

DT 26-OCT-1999 (first entry)
 XX

XX Hepatocyte growth factor hybrid protein Magic F-1 cDNA.

XX Magic Factor-1; human; hepatocyte growth factor; HGF; apoptosis;
 KW chemotherapy; toxicity; therapy; ss.

XX Homo sapiens.
 OS

XX Synthetic.
 OS

XX Key Location/Qualifiers
 FT sig_peptide 1..63
 FT /*tag= a
 FT mat_peptide 64..1725
 FT /*tag= b

XX WO9938967-A2.
 PN

XX 05-AUG-1999.
 PD

XX 27-JAN-1999; 99WO-EP00478.
 PF

XX 30-JAN-1998; 98IT-MI00179.
 PR

XX (DOMP-) DOMPE SPA.
 PA

XX Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;
 PI

XX WPI; 1999-494090/41.
 DR

XX P-PSDB; AAY06619.
 DR

XX Recombinant proteins from recombination of HGF and MSP structural
 PT domains, useful for protection against apoptosis induced by
 PT chemotherapeutics
 XX

PS Example 2; Page 56-57; 63pp; English.

XX This DNA sequence codes for Magic F-1 factor (see AAY06619), a
 CC recombinant protein composed of the hepatocyte growth factor (HGF)
 CC alpha chain signal peptide, hairpin loop and kringle domains 1 and
 CC 2 joined via a peptide linker to the HGF hairpin loop and kringle
 CC domains 1 and 2 and a polyhistidine tag. The portions of DNA
 CC encoding the various regions of Magic F-1 were obtained by PCR
 CC amplification and then recombined to obtain the hybrid sequence.
 CC Expression vectors, prokaryotic or eukaryotic host cells and a
 CC process for preparing recombinant proteins from HGF and MSP are
 CC claimed. Magic F-1 and other recombinant proteins comprising HGF
 CC and/or macrophage stimulating protein (MSP) structural domains are
 CC used to prevent or treat chemotherapeutic-induced toxicity such as
 CC myelotoxicity, hepatotoxicity, nephrotoxicity, mucotoxicity and
 CC neurotoxicity (claimed). They protect cells from death (apoptosis)
 CC induced by chemotherapy of, e.g. tumors. In particular, they can
 CC be used for expansion of marrow precursors, to increase
 CC proliferation of the haematopoietic precursors or to stimulate

CC their entry in the cycle. By modification of the proteolytic
 CC site, hybrid factors can be obtained which are activated by
 CC proteases of the endoplasmic reticulum (such as furines) during
 CC their synthesis. When the proteolytic site is removed, permanently
 CC immature forms of the factors can be obtained, having a potential
 CC partial agonistic or antagonistic activity. Different functional
 CC domains can be combined so as to modulate the biological effects.
 XX

SQ Sequence 1725 BP; 563 A; 399 C; 390 G; 373 T; 0 other;

Query Match 58.9%; Score 996.6; DB 20; Length 1725;

Best Local Similarity 83.2%; Pred. No. 1.7e-237;

Matches 1166; Conservative 0; Mismatches 214; Indels 21; Gaps 2;

Qy 292 tgtgacctctccagaagaagactacgtacggacctgcatcatgaacactgggttggg 351
 Db |||||||
 Qy 346 ttgacctctatgaacaaagactacattagaactgcatcatgtgtaaaagacagc 405
 Db |||||||
 Qy 352 taccggggcaccatggccacgacctgggtggcctgccaggttggagccacaag 411
 Db |||||||
 Qy 406 tacaagggaacagtatctatcactaagagtgcatacaaatgacacctggagttccatg 465
 Db |||||||
 Qy 412 ttccgaaatgatcaagaatcacgccccactctccggaatggcctggaaagaaactctgc 471
 Db |||||||
 Qy 466 ataccacacgaacacagcta-----tcggggtaaagacctacaggaataactctgt 516
 Db |||||||
 Qy 472 cgttaacctgatggcgaccccgagggtcctgtgtctacacacagacctgtgtgcgc 531
 Db |||||||
 Qy 517 cgaatctctcgggggaagaggggccctcgtgttccacaagaatccagagggcagc 576
 Db |||||||
 Qy 532 ttccagagctcggcatcaaatcctgcggggagggcgctgtgtctgtgcaatggcgag 591
 Db |||||||
 Qy 577 tacgaagtctgacattcctcagtttcagaagtgaatgcatgacacctgcaatggggag 636
 Db |||||||
 Qy 592 gaataccggcgctgtgacggcgacggagtcagggcgagtcgacacctggtggtatctt 651
 Db |||||||
 Qy 637 agttatcgaggtctcatggtatcagacaggaatcaggaagatttgcagcgtgggatcat 696
 Db |||||||
 Qy 652 cagacccgcacacacaccccttcagcggcggaagttcctcgaccaaagttcgtggagc 711
 Db |||||||
 Qy 697 cagacacacacccggcacaaattcttgcctgaaagatatcccgacaaaggcttctgat 756
 Db |||||||
 Qy 712 aactatgccgaatcctgacggctccgagcgccatggtgtacacctcagcatccgcag 771
 Db |||||||
 Qy 757 aattatgccgaatcctcagtcagggcgacggagccatggtgtactacttgcacctcac 816
 Db |||||||
 Qy 772 atcgagcgagagttctgtgacctccccctcgggtccgagggcagaccccgctctcgag 831
 Db |||||||
 Qy 817 acccgctggagtactgtgcaattaaacatgcgctgacaaagc-----ttcg 864
 Db |||||||
 Qy 832 ggcggtgcggttctggtggcggtgctccggtggtggtggttctctagaggacaaagg 891
 Db |||||||
 Qy 865 ggcggtggtggttctggtggcggtggtggtggtggtggtggtggtggtggtggtggt 924
 Db |||||||
 Qy 892 aaaaagaataataattcatgaattcaaaaaatcagcaaaactaccctaatcaataa 951
 Db |||||||
 Qy 925 aaaaagaataataattcatgaattcaaaaaatcagcaaaactaccctaatcaataa 984
 Db |||||||
 Qy 952 gatccagactgaagataaaaaaaccaaaagtgaatactgcagaccactgtgctaataga 1011
 Db |||||||
 Qy 985 gatccagactgaagataaaaaaaccaaaagtgaatactgcagaccactgtgctaataga 1044
 Db |||||||
 Qy 1012 tctactagaaataaaggacttccattcacttcaaggcttctgttttataaaagaaga 1071
 Db |||||||
 Qy 1045 tctactagaaataaaggacttccattcacttcaaggcttctgttttataaaagaaga 1104
 Db |||||||
 Qy 1072 aaacaatgctgtgttcccttcaatagcatgctcaagtggagtgaaaaagaattggc 1131
 Db |||||||
 Qy 1105 aaacaatgctgtgttcccttcaatagcatgctcaagtggagtgaaaaagaattggc 1164
 Db |||||||
 Qy 1132 catgaattgacctctatgaataaaagactacattagaactgcatcttggtaagga 1191
 Db |||||||

Db 661 caccagcacccttcgagccggaagtcttcctcgacccaaggtctggagcagacaactattgc 720
 Qy 721 cggaatctgacggtccgagcgccatggtgtctacactacggtatccgagatcgagcga 780
 Db 721 cggaatctgacggtccgagcgccatggtgtctacactacggtatccgagatcgagcga 780
 Qy 781 gaggttctgactcccccctgcggtccgagcgacagccccccctcgaggcc 834
 Db 781 gaggttctgactcccccctgcggtccgagcgacagcccccccaaggcc 834

RESULT 4

AAV7145
 ID AAV7145 standard; cDNA; 2216 BP.

XX AC AAT47145;

XX DT 27-MAR-1997 (first entry)

XX DE Macrophage stimulating protein C672X variant cDNA.

XX KW Macrophage stimulating protein; MSP; protein engineering;
 KW disulphide bond; ds.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT CDS 1..2136

FT sig_peptide /*tag= a

FT mat_peptide /*tag= b

FT misc_difference 2014..2016 /*tag= c

FT /*tag= d

FT /note= "nna at positions 2014-2016 represent an
 alternative codon to the native cysteine
 (tgc) codon, partic. GCN encoding alanine"

XX PN EP750040-A2.

XX PD 27-DEC-1996.

XX PF 18-JUN-1996; 96EP-0109750.

XX PR 20-JUN-1995; 95JP-0153309.

XX PA (TOYM) TOYO BOSEKI KK.

XX PA (TOYM) TOYOBO KK.

XX PI Hagiya M, Iwamoto J, Shimonishi M, Takehara T, Yoshikawa W;

XX WP1: 1997-044835/05.

XX DR P-PSDB; AAV07692.

XX PT Macrophage stimulating protein mutant - with cysteine deletion or
 PT substitution, can be produced recombinantly with minimal loss of
 PT activity due to incorrect disulphide bond formation

XX PS Claim 11; Page 12-14; 2lpp; English.

XX CC A DNA fragment (AAV47145) encodes a human macrophage stimulating
 CC protein (MSP) mutant (AAV07692) in which the native cysteine residue
 CC at position 672 is replaced by another amino acid, pref. alanine.
 CC The native gene is subjected to site-directed mutagenesis to alter
 CC codon 672. Alternatively, the Cys672 codon is deleted (see also
 CC AAV47149). The mutant MSPs can be mass produced in transfected host
 CC cells with minimal loss of activity due to incorrect disulphide
 CC bond formation.

XX SQ Sequence 2216 BP; 445 A; 665 C; 681 G; 422 T; 3 other;

Query Match 48.9%; Score 827.6; DB 18; Length 2216;
 Best Local Similarity 99.5%; Pred. No. 1.5e-195;
 Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 atgggggtgctcccaactcctgctctctgactcaatgcttaggggtcctcgagcgc 60
 Db 1 atgggggtgctcccaactcctgctctctgactcaatgcttaggggtcctcgagcgc 60
 Qy 61 tcgccattgaatgacttccaagtgtcccggtggcagagactacagcacctgtacatcg 120
 Db 61 tcgccattgaatgacttccaagtgtcccggtggcagagactacagcacctgtacatcg 120
 Qy 121 gtgtgcccggccttgccagagagatgtggcagatgtctgaagatgtgtgctgctgt 180
 Db 121 gtgtgcccggccttgccagagagatgtggcagatgtctgaagatgtgtgctgctgt 180
 Qy 181 gggcccttaatgactgcccggccttcactacacagctgagcagcattgtgccaactg 240
 Db 181 gggcccttaatgactgcccggccttcactacacagctgagcagcattgtgccaactg 240
 Qy 241 ctgccatgactcaaacactgcccacacaggtcggtctgtggcgctgtgacctc 300
 Db 241 ctgccatgactcaaacactgcccacacaggtcggtctgtggcgctgtgacctc 300
 Qy 301 ttccagaagaagactacgtacggaactgcataatgaacaaatgggtgtgggtaccgggc 360
 Db 301 ttccagaagaagactacgtacggaactgcataatgaacaaatgggtgtgggtaccgggc 360
 Qy 361 accatggcacgacgtgggtgctgctcctccagcttgagcagcagtcacgaat 420
 Db 361 accatggcacgacgtgggtgctgctcctccagcttgagcagcagtcacgaat 420
 Qy 421 gatcacagtagcacgccactctccggaatggcctggaagagaactcttgcgttaacct 480
 Db 421 gatcacagtagcacgccactctccggaatggcctggaagagaactcttgcgttaacct 480
 Qy 481 gatggcagcccgagggtctctgtgtacacaaacagccctgtgctgcttcagagc 540
 Db 481 gatggcagcccgagggtctctgtgtacacaaacagccctgtgctgcttcagagc 540
 Qy 541 tgcggcatcaaatcctgcggagcgcgtgtgtctgtgcaatggcaggaataccgc 600
 Db 541 tgcggcatcaaatcctgcggagcgcgtgtgtctgtgcaatggcaggaataccgc 600
 Qy 601 ggcggtgtagaccgcagagtcaggcgcgagtcagcgcgtgggtatcttcagccccg 660
 Db 601 ggcggtgtagaccgcagagtcaggcgcgagtcagcgcgtgggtatcttcagccccg 660
 Qy 661 caccagcacccttcgagccgggcaagtctcctgacccaaggtctgagcacaactatgc 720
 Db 661 caccagcacccttcgagccgggcaagtctcctgacccaaggtctgagcacaactatgc 720
 Qy 721 cggaatctgacggtccgagcgcgtgtgtctacactacagatccgagatcgagcga 780
 Db 721 cggaatctgacggtccgagcgcgtgtgtctacactacagatccgagatcgagcga 780
 Qy 781 gaggttctgacctccccctgcggtccgagcgacagccccctcgaggcc 834
 Db 781 gaggttctgacctccccctgcggtccgagcgacagccccccaaggcc 834

RESULT 5

AAV72083

ID AAV72083 standard; cDNA; 2216 BP.

XX AC AAV72083;

XX DT 12-APR-1999 (first entry)

XX DE Human MSP cDNA.

XX KW MSP; macrophage stimulating protein; apoptosis; human; treatment;

XX KW neuroendocrine cell; RON receptor; small cell lung carcinoma; tumour;

Db	561	agccggggcaagtctctcgaccgaagtcctgagcagacaactattcgcggaatcctgacggct	720
Qy	737	ccgagcgccatggtgtactacacacgagtcgcgacatcgagcagaggtctctgtgacctcc	796
Db	721	ccgagcgccatggtgtactacacgagtcgcgacatcgagcagaggtctctgtgacctcc	780
Qy	797	ccgcgtcggggtccgagggcagcccccgcctctcgagggc	834
Db	781	cccgctgctgggtccgagggcacagcccccgaagagggc	818
RESULT	11		
AAT62437			
ID	AAT62437	standard; cDNA to mRNA; 2219 BP.	
XX	AC		
XX	AC	AAT62437;	
DT	18-JUL-1997	(first entry)	
XX			
DE	XX	Human L5/3 partial clone #33 polymorphism #2.	
XX			
KW	KW	Human; growth factor; foetal; liver; probe; bovine; prothrombin; locus;	
KW	KW	polymorphism; transition; exon; intron; chromosome; kringle domain; ss;	
KW	KW	cell growth; tumour suppressor; hepatocyte growth factor; regeneration.	
XX			
OS	OS	Homo sapiens.	
XX			
PH	XX	Key	
FT	FT	Location/Qualifiers	
CDS		3..2120	
FT	FT	/*tag= a	
FT	FT	/product= part of human L5/3 protein	
FT	FT	3..77	
FT	FT	/*tag= b	
FT	FT	78..2117	
FT	FT	mat_peptide	
FT	FT	/*tag= c	
FT	FT	variation	
FT	FT	replace(618..620, TGC)	
FT	FT	/*tag= d	
XX			
PN	XX	US5606029-A.	
XX			
PD	25-FEB-1997.		
XX			
PF	14-MAY-1992;	92US-0882925.	
XX			
PR	14-MAY-1992;	92US-0882925.	
XX			
PR	18-JAN-1994;	94US-0184012.	
XX			
PA	(CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.		
XX			
PI	Degen SJ;		
XX			
XX	WPI: 1997-153621/14.		
DR	P-PSDB; AAW14267.		
XX			
PT	Human growth factor protein L5/3 - useful for altering cell growth,		
PT	e.g. as tumour suppressor		
XX			
PS	Disclosure; Column 17-24; 34pp; English.		
XX			
CC	This is the nucleotide sequence of the insert isolated from clone #33		
CC	and encodes part of a human growth factor designated L5/3 (AAW14270).		
CC	This sequence was isolated from a human foetal liver cDNA library using		
CC	a fragment of the corresponding genomic sequence as a probe. This		
CC	fragment was isolated from a human liver genomic DNA library using a 1200		
CC	bp probe derived from the bovine prothrombin gene. The longest human cDNA		
CC	isolated contained this sequence and a polymorphic sequence (AAT62436).		
CC	The sequences differ at nucleotides 619; T to G transversion. This clone		
CC	is not complete as the sequence starts 16 nucleotides downstream of the		
CC	initiator codon. The complete sequence has (AAT62440) has 18 exons, is		
CC	located at the D3F152 locus on human chromosome 3 and encodes for a		
CC	human growth factor which is an 80 kD single-chain protein containing 4		
CC	kringle domains. The protein can be used to alter cell growth (as a		
CC	growth factor or tumour suppressor) and has similar properties to the		

CC	hepatocyte growth factor that is actively involved in liver
CC	regeneration.
XX	
XX	
Sequence	2219 BP; 461 A; 660 C; 674 G; 420 T; 4 other;
Query Match	47.9%; Score 810; DB 18; Length 2219;
Best Local Similarity	99.4%; Pred. No. 3.4e-191;
Matches	813; Conservative 0; Mismatches 5; Indels 0; Gaps
Qy	17 tctgtgcttctgaactcaatgcttaagggttccttggcgagcgtcgccaattgaagtact 76
Db	1 tctgtgcttctgaactcaatctaggggttccttggcgagcgtcgccaattgaagtact 60
Qy	77 tccaagtgtctccgggacacagagctacagacacctgtcacatcggttgggtcccgggccctt 136
Db	61 tccaagtgtctccgggacacagagctacagacacctgtcacatcggttgggtcccgggccctt 120
Qy	137 ggcaggaggatgtgacagatgctgaagagtgtgtgctgctgtgggccccttaattggact 196
Db	121 ggcaggaggatgtgacagatgctgaagagtgtgtgctgctgtgggccccttaattggact 180
Qy	197 gccgggcttcacatacaacgctgagcagccatggttgccaaactgtctgcatagtgactcaac 256
Db	181 gccgggcttcacatacaacgctgagcagccatggttgccaaactgtctgcatagtgactcaac 240
Qy	257 actgcgccccacagagctcgacgttcttgggcgcgtgtgacctctccagagaagaagact 316
Db	241 actgcgccccacagagctcgacgttcttgggcgcgtgtgacctctccagagaagaagact 300
Qy	317 acgtacggacctgcatcatgaacaaatgggttgggtaccggggcaccatgtgccacgaccg 376
Db	301 acgtacggacctgcatcatgaacaaatgggttgggtaccggggcaccatgtgccacgaccg 360
Qy	377 tgggtggcctggccctgccaggcttggagcccaagtctcccgaaatgatcacaaatgatacgc 436
Db	361 tgggtggcctggccctgccaggcttggagcccaagtctcccgaaatgatcacaaatgatacgc 420
Qy	437 ccactctccggaatggcctggaagagaactcttcgcgttaacctgatggcgacccccggag 496
Db	421 ccactctccggaatggcctggaagagaactcttcgcgttaacctgatggcgacccccggag 480
Qy	497 gtcccttggctgctacacacagaccctgctgtgccttccagagctcgcgatcaaaatcct 556
Db	481 gtcccttggctgctacacacagaccctgctgtgccttccagagctcgcgatcaaaatcct 540
Qy	557 gccgggagggccgctgtgtctgtggtcaatggcagagaataaccgcgcggtgagaccgca 616
Db	541 gccgggagggccgctgtgtctgtggtcaatggcagagaataaccgcgcggtgagaccgca 600
Qy	617 cggagtcagggcgcgagctccagcgtctggatcttcagacaccgcaccagcacccttcg 676
Db	601 cggagtcagggcgcgagctccagcgtctggatcttcagacaccgcaccagcacccttcg 660
Qy	677 agccgcgcgaagtctctcgaccaagtcttgacgacaactattcccggaatccttgacggct 736
Db	661 agccgcgcgaagtctctcgaccaagtcttgacgacaactattcccggaatccttgacggct 720
Qy	737 ccgagcgcccatggtgtcacatactacgataccgcgagatcgagcgagaggttctgtgacctcc 796
Db	721 ccgagcgcccatggtgtcacatactacgataccgcgagatcgagcgagaggttctgtgacctcc 780
Qy	797 ccgcgtcggggtccgagggcacagcccgccctcgaggcc 834
Db	781 ccgcgtcggggtccgagggcacagcccgcccgcaagagccc 818
RESULT	12
AA62438	
ID	AA62438 standard; cDNA to mRNA; 2021 BP.
XX	
AC	
XX	AA62438;

(CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.

Degen SJ;

WPI; 1994-166645/20.

P-PSDB; AAR66599.

DNA from D3F15S2 locus of human chromosome 3 - encoding novel growth factor, L5/3, useful as probe for detecting pre-deposition towards cancer

Disclosure; Columns 23-28; 31pp; English.

The cDNA corresponding to a gene located at the D3F15S2 locus of human chromosome 3, designated L5/3, has been isolated (see AA079723). The gene codes for a protein composed of 4 kringle domains, followed by a serine protease-like domain. The hepatocyte growth factor has a similar structure but is only 50% identical to L5/3 at the amino acid sequence level. The L5/3 protein is useful for altering cell growth (as a growth factor or tumour suppressor). The L5/3 gene is also identical to the gene deleted from all human small lung cell carcinomas. The cDNA sequence of one L5/3 gene clone (i.e. clone #19, AA079725) had two parts of the coding region deleted when compared to the longest sequence AA079723. The region deleted included exon 13 and the 5' end of exon 18.

Sequence 2021 BP; 426 A; 610 C; 612 G; 373 T; 0 other;

Query Match 45.8%; Score 792.6; DB 15; Length 2021;

Best Local Similarity 99.5%; Pred. No. 6.8e-187;

Matches 795; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 36 atgttaggggtccctggcagcgtcccatggaatgacttccaaagtctccggggcag 95
Db 1 atgttaggggtccctggcagcgtcccatggaatgacttccaaagtctccggggcag 60
Qy 96 agagtacagacacctgtacatgogtgggtggtccggggtctggcaggaggtgtggcaga 155
Db 61 agagtacagacacctgtacatgogtgggtggtccggggtctggcaggaggtgtggcaga 120
Qy 156 tctgaagagtgctgtcgtctgtggtggtcccttaagtgaactcgcggccttcactacaa 215
Db 121 tctgaagagtgctgtcgtctgtggtggtcccttaagtgaactcgcggccttcactacaa 180
Qy 216 cgtgagcagccatggttgcgaactgtcctggaactcaacactgcgccacacagagct 275
Db 181 cgtgagcagccatggttgcgaactgtcctggaactcaacactgcgccacacagagct 240
Qy 276 gcggcgttctggcgtgtgaccttctccagaagaagactacgtacggaacctgcatcat 335
Db 241 gcggcgttctggcgtgtgaccttctccagaagaagactacgtacggaacctgcatcat 300
Qy 336 gaacaatgggttggtggtacgggggacacatgcccacagcgtggtggtccctgccatgca 395
Db 301 gaacaatgggttggtggtacgggggacacatgcccacagcgtggtggtccctgccatgca 360
Qy 396 ggcttggagccacaagtctccgaatgatcaaatgacacgcccactctccgaatggcct 455
Db 361 ggcttggagccacaagtctccgaatgatcaaatgacacgcccactctccgaatggcct 420
Qy 456 ggaagagaacttctccgttaacctgtatggcagccccggaggtccttgggtgtacacaaac 515
Db 421 ggaagagaacttctccgttaacctgtatggcagccccggaggtccttgggtgtacacaaac 480
Qy 516 agacctctgtgtcgtctccagagctgcgggacatacaaatcctgcgggagggcgcgtgtgt 575
Db 481 agacctctgtgtcgtctccagagctgcgggacatacaaatcctgcgggagggcgcgtgtgt 540
Qy 576 ctgggtgcaatggcaggaataaccgcgcgcgttagaccgcacgaggtcagggcgcgagtg 635
Db 541 ctgggtgcaatggcaggaataaccgcgcgcgttagaccgcacgaggtcagggcgcgagtg 600

Qy 636 ccagcgtctgggtatcttcagcaccgccagcaccctctgagccgggcaagtctctcga 695
Db 601 ccagcgtctgggtatcttcagcaccgccagcaccctctgagccgggcaagtctctcga 660
Qy 696 ccaagtctggagcacaactattgcccgaatcctgacggctccgagcgcacatggtgcta 755
Db 661 ccaagtctggagcacaactattgcccgaatcctgacggctccgagcgcacatggtgcta 720
Qy 756 cactcaggtatccgagatcagcagagaggttctgtgacctccccctgcgggtcccgaggc 815
Db 721 cactcaggtatccgagatcagcagagaggttctgtgacctccccctgcgggtcccgaggc 780
Qy 816 acagccccgcctcgagggc 834
Db 781 acagccccgcctcgagggc 799

RESULT 14

AAx87676

ID AAX87676 standard; cDNA; 2190 BP.

AC AAX87676;

DT 26-OCT-1999 (first entry)

DE HGF-MSP hybrid protein alphabeta-1 factor cDNA.

KW Alphabeta-1; human; hepatocyte growth factor; HGF; MSP;

KW macrophage stimulating protein; apoptosis; chemotherapy; toxicity; therapy; ss.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FH sig_peptide 1..63

FT mat_peptide /*tag= a

FT mat_peptide 63..2190

FT mat_peptide /*tag= b

XX W09938968-A1.

XX 05-AUG-1999.

XX 28-JAN-1999; 99WO-EP00502.

XX 30-JAN-1998; 98IT-MI00180.

XX (DOMP-) DOMPE SPA.

XX Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;

XX WPI; 1999-494091/41.

XX P-PSDB; AAY06621.

Recombinant proteins from recombination of HGF and MSP structural domains, useful for protecting cells against apoptosis induced by chemotherapeutics

Example 1; Page 52-53; 63pp; English.

This DNA sequence codes for alphabeta-1 factor (see AAY06621), a recombinant protein composed of the hepatocyte growth factor (HGF) alpha chain (i.e. the signal sequence, hairpin loop and kringle 1-4), the natural cleavage site of HGF, the macrophage stimulating protein (MSP) beta chain and a poly-histidine tag sequence. This structure allows the recombinant protein to interact with both the HGF receptor (Met) and the MSP receptor (Ron) and thereby induce biological responses which are synergistic and selective compared with the natural factor and truncated forms of the proteins. The portions of DNA encoding the various portions of alphabeta-1 are obtained by PCR amplification of HGF or MSP cDNA and then recombined to obtain the hybrid sequence. Expression vectors,

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 18:29:31 ; Search time 97.89 Seconds
(without alignments)
3272.187 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atgggggtgctccactctct.....accacaccaccactag 1692

Scoring table:
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	829.2	49.0	2232	1	US-08-334-177-1
2	829.2	49.0	2232	5	PCT-US95-13830-1
3	827.6	48.9	2216	2	US-08-666-082B-2
4	811.6	48.0	2219	1	US-07-882-925A-1
5	811.6	48.0	2219	1	US-08-184-012C-1
6	811.6	48.0	2262	1	US-07-882-925A-7
7	811.6	48.0	2262	1	US-08-184-012C-7
8	810	47.9	2219	1	US-07-882-925A-2
9	810	47.9	2219	1	US-08-184-012C-2
10	794.2	46.9	2021	1	US-07-882-925A-3
11	794.2	46.9	2021	1	US-08-184-012C-3
12	773	45.7	2289	1	US-07-838-410-2
13	771.4	45.6	2172	4	US-08-030-410-2
14	761	45.0	2288	1	US-08-290-937B-4
15	744.2	44.2	2184	1	US-07-815-333A-1
16	533	31.5	2188	1	US-07-882-925A-4
17	533	31.5	2188	1	US-08-184-012C-4
18	284.8	16.8	6100	1	US-07-882-925A-6
19	284.8	16.8	1284	3	US-08-184-012C-6
20	136.4	8.1	1284	3	US-08-985-526-24
21	136.2	8.0	645	3	US-08-985-526-22
22	136	8.0	2497	1	US-08-643-219-12
23	136	8.0	2497	2	US-09-131-995-12
24	136	8.0	2497	2	US-08-832-087B-12
25	136	8.0	2497	3	US-08-851-350-12
26	136	8.0	2497	4	US-09-132-154-12
27	136	8.0	2679	6	5200340-7

Patent No. 5200340

28	134.4	7.9	2753	1	US-07-854-603-1	Sequence 1, Appli
29	134	7.9	1134	4	US-09-206-059-29	Sequence 29, Appl
30	128.4	7.6	2296	1	US-07-750-080A-18	Sequence 18, Appl
31	128.4	7.6	2296	3	US-08-651-472-18	Sequence 18, Appl
32	114	6.7	6751	1	US-07-882-925A-5	Sequence 5, Appli
33	114	6.7	6751	1	US-08-184-012C-5	Sequence 5, Appli
34	60.4	3.6	1869	3	US-08-952-967-7	Sequence 7, Appli
35	60.2	3.6	4092	2	US-08-469-537A-106	Sequence 106, App
36	58	3.4	1947	1	US-07-998-972A-2	Sequence 2, Appli
37	58	3.4	1947	1	US-08-463-953-2	Sequence 2, Appli
38	58	3.4	1947	1	US-08-462-261-2	Sequence 2, Appli
39	58	3.4	1947	2	US-08-479-733A-24	Sequence 24, Appl
40	58	3.4	1947	3	US-08-487-427-24	Sequence 24, Appl
41	58	3.4	1947	3	US-08-479-727A-24	Sequence 24, Appl
42	58	3.4	1947	3	US-08-482-369A-24	Sequence 24, Appl
43	58	3.4	1947	5	PCT-US92-11357-2	Sequence 2, Appli
44	58	3.4	1947	5	PCT-US95-07439-24	Sequence 24, Appl
45	58	3.4	1988	1	US-07-750-080A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-334-177-1
; Sequence 1, Application US/08334177
; Patent No. 5696086
; GENERAL INFORMATION:
; APPLICANT: Avraham, Hava Karsenty
; APPLICANT: Godowski, Paul J.
; TITLE OF INVENTION: Methods and Kits Using Macrophage Stimulating Protein
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/334,177
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: 912
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2232 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-334-177-1

Query Match 49.0%; Score 829.2; DB 1; Length 2232;
Best Local Similarity 99.6%; Pred. No. 2.6e-211;
Matches 831; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 atgggggtgctccactctctctgactcaatgcttaggggtccctgggcagcgc 60

QY 661 caccagacccttcgagccgggaagttcttcgaccgaaggtctgacacacaactattgc 720
Db 671 CACCAAGACCCCTCGAGCGGGCAGATTCTTCGACCAAGGCTGGACGACAACTATTGC 730
QY 721 cggaaatctgacggctccgagcgccatgtgtctacactacgagatccgagatcagagca 780
Db 731 CGGAATCTGACGGCTCCGAGCGGCATGTGTCTACACTACGATCCGCGAGATCGAGCGA 790
QY 781 gaggttctgacctccccgcgtcggggtccgagggcacagccccgcctcagagggc 834
Db 791 GAGTCTGTGACTCCCGCGTCCGGGTCGAGGACAGACCCCGCCCAAGAGGCC 844

RESULT 3

US-08-666-082B-2

; Sequence 2, Application US/08666082B

; Patent No. 5916770

; GENERAL INFORMATION:

; APPLICANT: YOSHIKAWA, WATARU

; APPLICANT: SHIMONISHI, MANABU

; APPLICANT: IWAMOTO, JUNKO

; APPLICANT: TAKEHARA, TOYOHIRO

; APPLICANT: HAGIYA, MICHIO

; TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN VARIANT AND

; METHOD FOR PRODUCING THE SAME

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.

; STREET: TWO PRUDENTIAL PLAZA, SUITE 4900

; CITY: CHICAGO

; STATE: IL

; COUNTRY: US

; ZIP: 60601-6780

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/666,082B

; FILING DATE: 19-JUN-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 7-153309

; FILING DATE: 20-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: ROBERT F. GREEN

; REGISTRATION NUMBER: 27555

; REFERENCE/DOCKET NUMBER: 73843

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2216 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; ORIGINAL SOURCE:

; ORGANISM: Human

; FEATURE:

; NAME/KEY: sig peptide

; LOCATION: 1..54

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 1..2133

US-08-666-082B-2

Query Match 48.9%; Score 827.6; DB 2; Length 2216;

Best Local Similarity 99.5%; Pred. No. 6.8e-211;

Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 atgggggtggtccctcctctgtctgtactgaatgttaggggtccctcgtggcagcgc 60

|||||

Db 1 ATGGGTGGCTCCCACTCTCTGCTGCTTCTGACTCAATACCTTAGGGGTCCCTGGCAGCGC 60
QY 61 tcgccaatgaatgaacttccaaagtctccgggggcaagctacagacactctcatatgc 120
Db 61 TCGCCAATGAATGACTTCCAAAGTGTCTCCGGGGCACAGAGCTACAGCACCTCTCATATCGC 120
QY 121 gtgtgcccgggcttgagcagaggatgtgacagatctaaagatgtgtggtcgctgct 180
Db 121 GTGGTGCCTGGGCTTGGCAGGAGGATGTGGCAGATGCTGAAGAGTGTGTGGTTCGCTGT 180
QY 181 gggcccttaatgactgcgggcttccactacaacgtgagcagccatgggttgccaactg 240
Db 181 GGGCCCTTAATGACTGCCGGGCTTCCACTACAACTGACGACGCCATGGTGGCAACTG 240
QY 241 ctgcatggactcaaacactcgccccacacagagctgagggttcttgggcgtgtgacctc 300
Db 241 CTGCCATGGACTCAACACTCGCCCCACACGAGGCTCGGGCTTCTGGGCGCTGTGACCTC 300
QY 301 ttccagaagaagactacgtacggacctgcatcatgaataatgggttggttacccgggc 360
Db 301 TTCAGAAGAAGACTACGTACGACCTGTCATCATGAACATGGGGTGGGTACCGGGGC 360
QY 361 accatggccacagcctgggtgacctccctgcccagcttgagccacaagtcccgaat 420
Db 361 ACCATGGCCACGACCGGGTGGCTGCTCCCTGCCAGGCTTGGAGCCCAAGTTCCCGAAT 420
QY 421 gatcacaagtacacgcccactctccgggaatggcctggaagagaactctgcgtaacct 480
Db 421 GATCACAAGTACACGCCCACTCTCCGGAATGGCTTGGGAAGAACTTTCGCCGTAACCTT 480
QY 481 gatggcagccccggaggtccttggtgtctacacaacagacctgtgtgcttccagagc 540
Db 481 GATGGCAGCCCGAGGTCTTGGTGTCTACACAACAGACCCCTGTGTGCTTCCAGAGC 540
QY 541 tgcggcatcaaatcctcgggggagggccgctgtgtctgtgtcaatggcaggaataccgc 600
Db 541 TGCGGCATCAAAATCTCCCGGAGGCGCGTGTCTCTGGTCAATGGCGGGAATACCGC 600
QY 601 ggcgcggtagaccgcagcaggtcagggcgagtgccagcgctggatcttcagacccgc 660
Db 601 GCGCGGTAGACCGCAGGAGTACAGGGCGGCGAGTGCACGCGCTGGGATCTTCAGCACCG 660
QY 661 caccagacccccctcggcgggcaagtctcctcgaccacaaagtcttgagcagacaactatgc 720
Db 661 CACGACACCCCTTCGAGCGGGCAAGTTCCTCGACCAAGGCTCTGGACGCAACTATTGC 720
QY 721 cggaaatctgacggctccgagcgccatgtgtctacactacgagatccgacagatcgagca 780
Db 721 CGGAATCTCGCGCTCCGAGCGGCCATGTGTGTCTACACTACGGATCCCGACATCGAGCGA 780
QY 781 gagtctgtgacctccccgcgtcggggtccgagggcacagccccgcctcaggggc 834
Db 781 GAGTCTGTGACTCCCGCTCGGGGTCCGAGGCGACAGCCCCCGCCCAAGAGGCC 834

RESULT 4

US-07-882-925A-1

; Sequence 1, Application US/07882925A

; Patent No. 5315000

; GENERAL INFORMATION:

; APPLICANT: Degen, Sandra J. F.

; TITLE OF INVENTION: Gene for a growth factor and its cDNA and

; protein

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Gregory Lunn

; STREET: Wood, Herron & Evans, 2700 Carew Tower

; CITY: Cincinnati

; STATE: Ohio

; COUNTRY: USA

; ZIP: 45202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb

TISSUE TYPE: liver;
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #33
POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/D3F152
FEATURE:
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: Includes five polymorphisms at the nucleotide level; one of which results in an amino acid substitution (nucleotide 619). Sequence ID NO:2: contains the identical sequence with Patent No. 560629
OTHER INFORMATION: polymorphic amino acid.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 2219
US-08-184-012C-1

Query Match 48.0%; Score 811.6; DB 1; Length 2219;
Best Local Similarity 99.5%; Pred. No. 1.2e-206;
Matches 814; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 17 tctgtctcttctgactcaatgcttaggggtccctgggcagcgctcgccattgaatgact 76
Db 1 tctgtctcttctgactcaatgcttaggggtccctgggcagcgctcgccattgaatgact 60
Qy 77 tccaaagtctccggggcacagagctacagcacctgctacatgctggtggtcccggtctt 136
Db 61 tccaaagtctccggggcacagagctacagcacctgctacatgctggtggtcccggtctt 120
Qy 137 ggcaggagatgtggcagatgctgaagagtgtgctggtgctgtggtggcccttaatggact 196
Db 121 ggcaggagatgtggcagatgctgaagagtgtgctggtgctgtggtggcccttaatggact 180
Qy 197 gccggccttcactacaaatgagcagcagcagcagcagcagcagcagcagcagcagcagc 256
Db 181 gccggccttcactacaaatgagcagcagcagcagcagcagcagcagcagcagcagcagc 240
Qy 257 actgcggccacacagagctgcgggttctggcgtgctgacgtcttcagagaagaagact 316
Db 241 actgcggccacacagagctgcgggttctggcgtgctgacgtcttcagagaagaagact 300
Qy 317 acgtacgacctgcatcatgaacatgggttgggttacccggggcaccatggccacgacgg 376
Db 301 acgtacgacctgcatcatgaacatgggttgggttacccggggcaccatggccacgacgg 360
Qy 377 tgggtggcctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 436
Db 361 tgggtggcctccctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 420
Qy 437 ccaactccggaatggcctggaagaaactctgcgttaaccctgatggcaccgccggag 496
Db 421 ccaactccggaatggcctggaagaaactctgcgttaaccctgatggcaccgccggag 480
Qy 497 gtccttggtgtacacaaagaccctgctgctgcttcacagagctgcggatcaaatctt 556
Db 481 gtccttggtgtgtacacaaagaccctgctgctgcttcacagagctgcggatcaaatctt 540
Qy 557 gccgggagggcgctgtgtctgtgtaagtgcgaggaataccgcggcggttagaccgca 616
Db 541 gccgggagggcgctgtgtctgtgtaagtgcgaggaataccgcggcggttagaccgca 600
Qy 617 cggagtcaggcgagtgccagcgtgggatcttcagcaccgcaccagcaccctctcg 676
Db 601 cggagtcaggcgagtgccagcgtgggatcttcagcaccgcaccagcaccctctcg 660
Qy 677 agccggggaagtctctcgaccaaaggtctggacgacaaactatttcgggaatcctgacggct 736
Db 661 agccggggaagtctctcgaccaaaggtctggacgacaaactatttcgggaatcctgacggct 720
Qy 737 ccgagcgccatggtgtctactacagatccgcagagatcgagcagaggttctgtgacctcc 796
Db 721 ccgagcgccatggtgtctactacagatccgcagagatcgagcagaggttctgtgacctcc 780

Qy 797 ccgctgctgggtccgagggcacagcccccgcctcgagggc 834
Db 781 CCCGCTCGGGTCCGAGGCACAGCCCGCCCAAGAGGCC 818
RESULT 6
US-07-882-925A-7
Sequence 7, Application US/07882925A
Patent No. 5315000
GENERAL INFORMATION:
APPLICANT: Degen, Sandra J. F.
TITLE OF INVENTION: Gene for a growth factor and its cDNA and
TITLE OF INVENTION: protein
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSER: Gregory Lunn
STREET: Wood, Herron & Evans, 2700 Carew Tower
CITY: Cincinnati
STATE: Ohio
COUNTRY: USA
ZIP: 45202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 800 kb
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.3
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/882,925A
FILING DATE: 19920514
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Gregory
REGISTRATION NUMBER: 29,945
REFERENCE/DOCKET NUMBER: CMC 57
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 241-2324
TELEFAX: (513) 421-7269
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2262 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: Identical to sequence ID NO: 1: with 5' and 3'
DESCRIPTION: adaptors added to make a full-length cDNA
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #33 including 5' and 3' adaptors
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 2262
US-07-882-925A-7

Query Match 48.0%; Score 811.6; DB 1; Length 2262;
Best Local Similarity 99.5%; Pred. No. 1.3e-206;
Matches 814; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 17 tctgtgtcttctgactcaatgcttaggggtccctgggcagcgctcgccattgaatgact 76
Db 32 tctgtgtcttctgactcaatgcttaggggtccctgggcagcgctcgccattgaatgact 91
Qy 77 tccaaagtctccggggcacagagctacagcacctgctacatgctggtggtcccggtctt 136
Db 92 tccaaagtctccggggcacagagctacagcacctgctacatgctggtggtcccggtctt 151
Qy 137 ggcaggagatgtggcagatgctgaagagtgtgctggtcgctgtggcccttaatggact 196

Qy 617 cggagtcaggcgagtcgagcgcgtggtatcttcacacccgcacccagacccttcg 676
Db 632 CGGAGTCAGGCGCGAGTGCAGCGCTGGGATCTTCAGACCGCCGACCCAGCCCTTCG 691
Qy 677 agccgggcaagtctctgaccaaagtctggacgacaactattgccgggaatcctgacggct 736
Db 692 AGCCGGGCAAGTCTTCACCAAGGCTGGACGACAATATTGCCGGAATCCTGACGGCT 751
Qy 737 ccgagcgccatggtgctacactacggatccgcagatcgagcgagagtgcttgacctcc 796
Db 752 CCGAGCGGCATGCTGCTACACTACGGATCCGCAGATCGAGCGAGAGTCTCTGACCTCC 811
Qy 797 ccgctgcgggtccgagggcacagccccccctcgagggc 834
Db 812 CCCGCTGCGGTCGAGGCACAGCCCCCGCCCAAGAGGCC 849

RESULT 8

US-07-882-925A-2
; Sequence 2, Application US/07882925A
; Patent No. 5315000
; GENERAL INFORMATION:
; APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gregory Lunn
; STREET: Wood, Herron & Evans, 2700 Carew Tower
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 kb
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.3
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/882,925A
; FILING DATE: 19920514
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Gregory
; REGISTRATION NUMBER: 29,945
; REFERENCE/DOCKET NUMBER: CMC 57
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 241-2324
; TELEFAX: (513) 421-7269
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2219 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; DEVELOPMENTAL STAGE: fetal
; TISSUE TYPE: liver
; IMMEDIATE SOURCE:
; LIBRARY: cDNA
; CLONE: #33
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: human 3p21/D3F15S2
; FEATURE:
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: Includes five polymorphisms at the nucleotide
; OTHER INFORMATION: level; one of which results in an amino acid substitution (no
; OTHER INFORMATION: 619). Sequence ID NO:1: contains the identical sequence with
; OTHER INFORMATION: polymorphic amino acid.
; INFORMATION:

RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 2219
US-07-882-925A-2
Query Match 47.9%; Score 810; DB 1; Length 2219;
Best Local Similarity 99.4%; Pred. No. 3.3e-206;
Matches 813; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 17 tctgtgtcttctgactcaatggttaggggtccctggcgagcgtctgcccattgaatgact 76
Db 1 TCTGTGCTCTTCTGACTCAATACTTAGGGGTCCCTGGCGAGCGCTGCCATTGAATGACT 60
Qy 77 tccaagtgtctcggggcacagagctacagcacctgtacatgcggtgtgtcccggggctt 136
Db 61 TCCAAGTGTCTCGGGGCACAGAGCTACAGCACCTGTACATGCGGTGGTCCCGGGGCTT 120
Qy 137 ggcagagagatgtggcagatgctgaagagtgtgctgctgctgtggcccttaatgact 196
Db 121 GGCAGGAGGATGTGGCAGATGCTGAAGAGTGTGCTGCTGTGGCCCTTAATGACT 180
Qy 197 gcccggccttccactacaactgagcagccatggttgcgaactgctgcccattgactcaac 256
Db 181 GCCGGGCTTCCACTACAACGTGAGCAGCCATGTTGCCAACTGCTGCCATGGACTCAAC 240
Qy 257 actcgcccccacacagagctgcgcttctggtggtgctgacacctcttccagaagaagact 316
Db 241 ACTCGCCCCACACAGAGGCTCGCGGCTTCTGGGCGCTGTGACCTCTTCCAGAAGAAGACT 300
Qy 317 agtctcgacctcatcatgaacaatgggttgggtaccggggcaccatgcccacgaccg 376
Db 301 AGTACGGACCTGCATCATGAACAATGGGGTTGGGTACCGGGGCACCATGGCCACGACCG 360
Qy 377 tgggtggcctgccctgccaggcttgagccacaagtctcccgaaatgatcacaagtacacgc 436
Db 361 TGGGTGGCTGCCTGCCAGGCTTGGAGCCACAAGTTCGCCGAATGATCACAAGTACACGC 420
Qy 437 caactctcgggaatgacctggaagaacttctgctgaacctctgacccctgacggaccccgag 496
Db 421 CCACCTCTCCGGAATGGCTTGGAGAGAACTTCTGCCGTAACCTGTATGGGACCCCCGGAG 480
Qy 497 gtccttgggtgtacacaacagacccctgctgtgccttccagagctgcggcatcaaatcct 556
Db 481 GTCTTGGTGTACACAACACAGACCCCTGCTGTGCTTCCAGAGCTGCGGCATCAATCCT 540
Qy 557 gcccggagggcgctgtgtctgtgcaatggcgagggaatacccgcgcggtagacgcga 616
Db 541 GCGGGAGGCGCGCTGTGTCTGTGCAATGGCGAGGAATACCGCGCGCGGTAGACCGCA 600
Qy 617 gggagtcaggcgcgagtgccagcgtggatcttccagcccccagcaccaccccttcg 676
Db 601 CGGAGTCAGGCGCGAGTTCACGCTGGGATCTTCAGCACCCCGCACAGACCCCTTCG 660
Qy 677 agccgggcaagtctcgcaccaaaggctcgcagcagcagcagcagcagcagcagcagcagc 736
Db 661 AGCCGGGCAAGTCTTCGACCAAGGCTGGACGACAATATTGCCGGAATCTCTGACGGCT 720
Qy 737 ccgagcgccatggtgctacactacgagtcgcagatccgcagatcgagcgagagtgcttgacctcc 796
Db 721 CCGAGCGGCATGCTGCTACACTACGGATCCGCAGATCGAGCGAGAGTCTCTGACCTCC 780
Qy 797 ccgctgcgggtccgagggcacagccccccctcgagggc 834
Db 781 CCCGCTGCGGTCGAGGCACAGCCCCCGCCCAAGAGGCC 818

RESULT 9
US-08-184-012C-2
; Sequence 2, Application US/08184012C
; Patent No. 5606029
; GENERAL INFORMATION:
; APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: protein

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 2021 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #19

POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/D3F15S2
FEATURE:

IDENTIFICATION METHOD: experimental
OTHER INFORMATION: This sequence is a variant where two regions
OTHER INFORMATION: were found to be deleted when compared to SEQ ID NO:1.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 2021
US-07-882-925A-3

Query Match 46.9%; Score 794.2; DB 1; Length 2021;
Best Local Similarity 99.6%; Pred. No. 5.1e-202;
Matches 796; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 36 atgcttaggggtccctgggagcgctgcgcattgaattgaattccaaagtccaaagtccggggcacc 95
DB 1 ATGCTTAGGGGTCCCTGGGAGCGCTCGCATTTGAATGAATGACTTCCAAAGTCTCCGGGGCAC 60
QY 96 agagctacagacactgtacatcggtgggtgcccggcccttgccagagatgtggcaga 155
DB 61 AGAGCTACAGACCTGCTACATCGGTGGTGGCCGGCCCTTGGCAGGAGGATGTGGCAGA 120
QY 156 tgctgaagagtgctgtgctgtggtggcccttaattgaactgcccggccctccactacaa 215
DB 121 TGCTGAAGAGTGTGCTGCTGCTGGGCCCTTAATGACTGCGGGCCCTTCCACTACAA 180
QY 216 cgtgagcagccatggttgccaaactgtgcccactgaactcaactcgccacacagagct 275
DB 181 CGTGAGCAGCATGTTGGCCAACTGCTGCCATGGACTCAACACTCGCCGCCACACAGGCT 240
QY 276 gcggcgttcggcgctgtgacctcttcagaagaagactagactagcagacctgcatcat 335
DB 241 CGCGCGTCTGGGCGCTGTGACCTCTTCCAGAAAGAAAGACTACGTACGGACCTGCATCAT 300
QY 336 gaacaaatgggttgggtaccgggggaccatggccacgacctggtggtgacctgacctgcca 395
DB 301 GAACAATGGGGTGGGTACCGGGGCACCATGTGGCCACACCGCTGGGTGGCTGCCCTGCCA 360
QY 396 ggcttgagccacaaagtctccgaatgatcacaaagtacacacccactctccggaatggcct 455
DB 361 GGCTTGAGGACCAAGTTCCTCGAATGATACAACTACACGCCACTCTCCGGAATGGCT 420
QY 456 ggaagagaactctgcgtaacctgtagcagcccgagggcttcttgggtgctacacaaac 515
DB 421 GGAAGAGAAGTCTTCGCGGTAAACCTGTATGGACACCGCGAGGTCTTGGTGTCTACAAAC 480
QY 516 agacctgtgtgctgtccagagctcagcatcaataatcctgcggagccgctgtgt 575
DB 481 AGACCCCTGCTGTGGCTTCCAGAGCTCGGCATCAATCTTCCGGGAGGCGCGGTGTGT 540
QY 576 ctggtgcaatggcgaggaataaccgcgcggttagacccgacgagtcagggcgcgagtg 635
DB 541 CTGGTGAATGGCGAGGAATACCGCGCGGTAGACCGCAGGAGTCAGGCGCGGAGTG 600
QY 636 ccagcgtggatcttcagacacccgacacgaccccttcagccggttccttcga 695
DB 601 CCAGCGCTGGGATCTTCAGCACCCGACACCGACCCCTTCGAGCGGGCAAGTTCCTCGA 660

QY 596 ccaaggtctgacgacaaactattgcccgaatactcctacgggctccgagcgccatggtgcta 755
DB 661 CCAAGGTCTGGACGACAACCTATTGCCGAATCTGACGGCTCCGAGCGGCCATGTTGCTA 720
QY 756 cactacagatccgagatcgagcagagttctgtgtaacctccccccctcggtcggtccgagcg 815
DB 721 CACTACGATCCGAGATCGAGATCGAGGAGTTCTGTGACCTCCCCCGCTCGGGTCCGAGGC 780
QY 816 acagcccgctcctcgagggc 834
DB 781 ACAGCCCGCCCAAGAGGCC 799

RESULT 11

US-08-184-012C-3
Sequence 3, Application US/08184012C
Patent No. 5606029
GENERAL INFORMATION:
APPLICANT: Degen, Sandra J. F.
TITLE OF INVENTION: Gene for a growth factor and its cDNA and
TITLE OF INVENTION: protein
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gregory Lunn
STREET: Wood, Herron & Evans, 2700 Carew Tower
CITY: Cincinnati
STATE: Ohio
COUNTRY: USA
ZIP: 45202
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.2
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/184,012C
FILING DATE: 1/18/94
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Lund, Gregory
REGISTRATION NUMBER: 29,945
REFERENCE/DOCKET NUMBER: CMC 57
TELECOMMUNICATION INFORMATION:
TELEPHONE: (513) 241-2324
TELEFAX: (513) 421-7269
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2021 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
DEVELOPMENTAL STAGE: fetal
TISSUE TYPE: liver
IMMEDIATE SOURCE:
LIBRARY: cDNA
CLONE: #19
POSITION IN GENOME:
CHROMOSOME/SEGMENT: human 3p21/D3F15S2
FEATURE:
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: This sequence is a variant where two regions
OTHER INFORMATION: were found to be deleted when compared to SEQ ID NO:1.
PUBLICATION INFORMATION:
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 2021
US-08-184-012C-3

Query Match 46.9%; Score 794.2; DB 1; Length 2021;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 18:27:22 ; Search time 2582.57 Seconds
(without alignments)
6193.144 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atgggggtgtccaccctct.....accacaccaccaccactag 1692

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_est3.*
- 4: gb_est4.*
- 5: gb_est5.*
- 6: gb_est6.*
- 7: gb_est7.*
- 8: gb_est8.*
- 9: gb_est9.*
- 10: gb_est10.*
- 11: gb_est11.*
- 12: gb_est12.*
- 13: gb_est13.*
- 14: gb_est14.*
- 15: gb_est15.*
- 16: gb_est16.*
- 17: gb_est17.*
- 18: gb_est18.*
- 19: gb_est19.*
- 20: gb_est20.*
- 21: gb_est21.*
- 22: gb_est22.*
- 23: gb_est23.*
- 24: gb_est24.*
- 25: gb_est33.*
- 26: gb_est34.*
- 27: gb_est35.*
- 28: gb_est36.*
- 29: gb_est37.*
- 30: gb_est38.*
- 31: gb_est39.*
- 32: gb_est40.*
- 33: gb_estba.*
- 34: em_estfun.*
- 35: em_esthum1.*
- 36: em_esthum2.*
- 37: em_esthum3.*
- 38: em_esthum4.*
- 39: em_esthum5.*
- 40: em_esthum6.*
- 41: em_esthum7.*
- 42: em_esthum8.*
- 43: em_esthum9.*

- 44: em_esthum10.*
- 45: em_esthum11.*
- 46: em_esthum12.*
- 47: em_esthum13.*
- 48: em_esthum14.*
- 49: em_esthum15.*
- 50: em_esthum16.*
- 51: em_esthum17.*
- 52: em_esthum18.*
- 53: em_esthum19.*
- 54: em_esthum20.*
- 55: em_esthum21.*
- 56: em_esthum22.*
- 57: em_esthum23.*
- 58: em_esthum24.*
- 59: em_esthum25.*
- 60: em_esthum26.*
- 61: em_esthum27.*
- 62: em_esthum28.*
- 63: em_estin1.*
- 64: em_estin2.*
- 65: em_estin3.*
- 66: em_estin4.*
- 67: em_estin5.*
- 68: em_estom1.*
- 69: em_estom2.*
- 70: em_estov1.*
- 71: em_estov2.*
- 72: em_estpl1.*
- 73: em_estpl2.*
- 74: em_estpl3.*
- 75: em_estpl4.*
- 76: em_estpl5.*
- 77: em_estpl6.*
- 78: em_estpl7.*
- 79: em_estpl8.*
- 80: em_estpl9.*
- 81: em_estpl10.*
- 82: em_estro1.*
- 83: em_estro2.*
- 84: em_estro3.*
- 85: em_estro4.*
- 86: em_estro5.*
- 87: em_estro6.*
- 88: em_estro7.*
- 89: em_estro8.*
- 90: em_estro9.*
- 91: em_estro10.*
- 92: em_estro11.*
- 93: em_estro12.*
- 94: em_estro13.*
- 95: em_estro14.*
- 96: em_estro15.*
- 97: em_estro16.*
- 98: em_estro17.*
- 99: em_estro18.*
- 100: em_estro19.*
- 101: em_estro20.*
- 102: gb_est25.*
- 103: gb_est26.*
- 104: gb_est27.*
- 105: gb_est28.*
- 106: gb_est29.*
- 107: gb_est30.*
- 108: gb_est31.*
- 109: gb_est32.*
- 110: gb_est41.*
- 111: gb_est42.*
- 112: gb_est43.*
- 113: gb_est44.*
- 114: gb_est45.*
- 115: gb_est46.*
- 116: gb_est47.*

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description
1	736	43.5	975	106	AL546558	AL546558
2	707	41.8	919	106	AL549164	AL549164
3	594.4	35.1	872	106	AL550803	AL550803
4	526.2	31.1	828	106	AL546883	AL546883
5	475.2	28.1	766	148	BF383652	BF383652
6	456	27.0	889	106	AL558014	AL558014
7	455.8	26.9	780	146	BF234001	BF234001
8	446	26.4	734	148	BF384148	BF384148
9	428.4	25.3	652	148	BF385836	BF385836
10	414.2	24.5	844	106	AL546856	AL546856
11	405.4	24.0	543	139	BE723534	BE723534
c	389.2	23.0	899	106	AL573472	AL573472
12	350.2	20.7	506	1	AA037738	AA037738
13	349.6	20.7	518	139	BE755132	BE755132
14	230.2	13.6	263	121	AW815873	AW815873
15	207.4	12.3	574	165	BE234370	BE234370
16	184.4	10.9	412	110	AV759905	AV759905
17	175.6	10.4	288	161	BB598484	BB598484
c	166	9.8	484	1	AA037786	AA037786
19	153.2	9.1	661	112	AW175268	AW175268
20	145.6	8.6	830	106	AL534044	AL534044
c	135.8	8.0	753	227	AQ381428	AQ381428
23	131.6	7.8	778	18	AI317783	AI317783
24	131.4	7.8	565	144	BF079866	BF079866
25	128	7.6	768	18	AI317745	AI317745
26	125.8	7.4	837	16	AI118965	AI118965
27	124.8	7.4	844	21	AI530175	AI530175
28	122.8	7.3	470	232	AQ707814	AQ707814
29	122	7.2	220	161	BB587413	BB587413
c	122	7.2	470	166	BB326846	BB326846
31	121.2	7.2	780	18	AI317794	AI317794
32	118.4	7.0	914	106	AL531543	AL531543
33	117.4	6.9	909	106	AL557971	AL557971
34	115.4	6.8	980	17	AI226536	AI226536
35	112.6	6.7	788	16	AI121047	AI121047
36	112.2	6.6	819	16	AI114957	AI114957
37	111.8	6.6	817	24	AI746861	AI746861
38	110.8	6.5	773	16	AI119053	AI119053
39	109.8	6.5	897	111	AW106317	AW106317
40	108	6.4	918	23	AI663558	AI663558
41	107.4	6.3	752	146	BF235733	BF235733
42	106.8	6.3	842	16	AI118969	AI118969
43	106.4	6.3	265	137	BE576686	BE576686
44	105	6.2	910	21	AI530076	AI530076
45	101.4	6.0	794	24	AI746972	AI746972

ALIGNMENTS

RESULT 1
AL546558
LOCUS AL546558 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI030YH12 5
DEFINITION prime, mRNA sequence. 16-FEB-2001
ACCESSION AL546558.1 GI:12879791
VERSION EST.
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 975)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 975
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI030YH12"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 311 a 228 c 215 g 220 t
ORIGIN
Query Match 43.5%; Score 736; DB 106; Length 975;
Best Local Similarity 96.8%; Pred. No. 4.2e-189; Gaps 2;
Matches 778; Conservative 0; Mismatches 10; Indels 16;
Qy 874 tctctagagggacaaagaaagaaatacaatcatgaattcaaaatacagcaag 933
Db 87 TATGCAGAGGGACAAAGGAAAGAGAAATACATTCATGATTCATAAATCAGCAAG 146
Qy 934 actaccctaatcaaatagatccagcactgaagataaaacacaaaggaatactgca 993
Db 147 ACTACCCATCAATCAATAGATCCAGCAGTGAAGATAAAACCAAAAGTGAATCTGCA 206
Qy 994 gaccatgtgctaagatgtagtaagataaaggactccatctcactgcaaggctttt 1053
Db 207 GACCAATGTCCTAATAGATGTAAGAGTAAAGAGCTTCCATTCCTGCAAGGCTTTT 266
Qy 1054 gttttgtataagcaagaaacaatgcctctgttcccttcaatagcatgcaagtgga 1113
Db 267 GTTTTGTATAAGCAAGAAACAATGCCCTGTGTTCCCTTCAATAGCATGTCAAGTGA 326
Qy 1114 gtgaaaaagaattggccatgaattgacctctatgaaacaaagactacattagaac 1173
Db 327 GTGAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAAACAAAGACTACATAGAAC 386
Qy 1174 tgcatcattgttaaaggacgagctacaaagggaacagtagtattctactaagagtgcatc 1233
Db 387 TGCATCATTTGGTAAAGGACGAGCTACAAGGGAACAGTATCTATCATAAGAGTGGCATC 446
Qy 1234 aaatgcagccctgagttccatgtatccacacagcaacac-----agctat 1278
Db 447 AAATGTCAGCCCTGGAGTTCCATGTATACCACACAGACAGCTTTTTCCTTCAGCAT 506
Qy 1279 cggggtaaagacacagaaactactctcgaaatccctcgagggaagaaagggagccc 1338
Db 507 CGGGGTAAAGACCTACAGGAAACTACTGTGAAATCTCGAGGGGAAAGAGGGGACCC 566
Qy 1339 tgggtttcacagcaatccagaggtacgctacgaagctgtgacattcctcagtggtca 1398
Db 567 TGGTGTTCACAAAGCAATCCAGAGTACCTAGGAAGTCTGTGACATTCCTCAGTGTCA 626
Qy 1399 gaagtgaatgcacgtgcaatggggagagttatcgaggtctcgtggtatcgatcacagaa 1458
Db 627 GAAGTTGAATGATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCATACAGAA 686
Qy 1459 tcaggcaagattgtcagcgctgggagatcatcagacacacacccgacacaaattctgcct 1518
Db 687 TCAGGCAAGATTGTGACGGCTGGGATCATCAGACACACACCGGCACAAATTTCTTGCCT 746
Qy 1519 gaaagatatcccgacaaagggtttgtgatataattgtccgcaatcccgatggccagccg 1578

```

BASE COUNT 274 a 212 c 110 g 266 t 10 others
ORIGIN

```

Query Match 35.1%; Score 594.4; DB 106; Length 872;
Best Local Similarity 89.8%; Pred. No. 1.3e-150;
Matches 666; Conservative 9; Mismatches 50; Indels 17; Gaps 3;

Qy	874	tctctagagggacaaggaagaaagaatacaattcatgaattcaaaaaatcagccaag	933
Db	132	TATTTCATATTGACAAATTAATAATAATAACAATTCAATTAATTCAAAAAATCATCAAT	191
Qy	934	actacccttaatacaaaaatagatccagcactgaagataaaaaacccaaaaagtgaaatactgcga	993
Db	192	ACTACCCCTAATCAAAATATATCCATCACTTAATATAAAACCAAAAAATTTAATTACTGCA	251
Qy	994	gaccaattgctaaatagatgtactaggaaagagcattccattcacttgcgaaggtctttt	1053
Db	252	TACCAATTTTCTAATAATATTTACTATTATAAATTAATTCCTCCATTCATTCCAANKCTTTT	311
Qy	1054	gtttttgataaagcaagaacaacatgcctctgtgttcccttccaatagatcgaatgcga	1113
Db	312	TTTTTTTATAAATCAAGAAACAATTCCTCTGTGTTCCTTCAATGAGCATTTCAAGTTGA	371
Qy	1114	gtgaaaaagaatttggccatgaatttgacctctctatgaaaaacaagactacattagaaac	1173
Db	372	GTTTAAAAAATAATTTTGCCATTAAATTTTACCTCTATTATAAAACAAGAGACTACATTTAKAAAC	431
Qy	1174	tgcattcatgtttaaaggccgacgtcacagggaacagtatctatcactaagaatgagcctc	1233
Db	432	TTTCAATCATYGTAAAGGACGACGTACAAGGGAACAGTATCTATCACTTAAGAGTGGCATC	491
Qy	1234	aaatgtcacccctggagttccatgataccacacgaacac-----agctat	1278
Db	492	AAATGTGAGCCCTGGAGTTCCATGATACACACATAACACATCTTTTTCCTTCTATCTAT	551
Qy	1279	cggggttaagacctacaggaataactactgtcgaaatcctcgagggggaagaggggagccc	1338
Db	552	CTTGGTAAAGACCTACAGGAAACTACTGTCCGAATCTCTAYGTGAATAMGKGTGA-CC	610
Qy	1339	tgggtttccacaagcaatccagaggtacgctacgaagtctgt-gacattcctcagtgctc	1397
Db	611	TGGTGYTTTACAAGCAATCCAGAGGTAGCTACGNAGTCTGTGGACATTTCTTCAAGTGTTC	670
Qy	1398	agaagttgaatgcatagcctgcaatggggagagttatcgaggtctcatgatcatacaga	1457
Db	671	AGAAATTGAATGCATGACCTGCAATGGGGAGAGTTTATCGAGGTCTCATGGATCATACAGA	730
Qy	1458	atcaggcaagatttgcagcgtgggatcatcagacacccacacccgggacaaaattcttgcc	1517
Db	731	ATCAGGCAAGATTTGTTCAGCGCTGGGATTCATCAGACACACACCGGCACAAATTTCTTGGC	790
Qy	1518	tgaagatatcccgacaagggttttgatgataattattgcgcgaatcccgatggccagcc	1577
Db	791	TGAAGAATATCCGCACAGGGCTTTTGATGATAATTTATTTGCCGCCAATCCCGATGGCCAGCS	850
Qy	1578	gaggccatgggtgctatactctt	1599
Db	851	GAGGCCATGGTGCTATACTCTTT	872

RESULT	4	
AL546883		
LOCUS		16-FEB-2001
DEFINITION	AL546883 828 bp mRNA EST	
	AL546883 LIT_NFL006_PL2 Homo sapiens cDNA clone CSODI026YD16 5	
	prime, mRNA sequence.	
ACCESSION	AL546883	
VERSION	AL546883.1	
KEYWORDS	GI:12880433	
SOURCE	EST.	
	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	1 (bases 1 to 828)
TITLE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.: Full-length cDNA libraries and normalization
COMMENT	Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES	Location/Qualifiers
source	1..828

WORKS sources

1. .828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS01026YD16"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : liang@lifetech.com URL : <http://fulllength.invitrogen.com>
290 a 183 c 157 g 196 t 2 others

Query Match	31.1%	Score 526.2	DB 106	Length 828
Best Local Similarity	98.7%	Pred. No. 4.6e-132		
Matches 528	Conservative 2	Mismatches 5	Indels 0	Gaps 0
Qy 874	tcttagaggcacaagaagaagaatacaattcatgaattcaaaaatcagcaag	933		
Db 96	TATCAGAGGACAAAGAAAGAAATACATTTCTGAATTCAAAATATCAGCAAG	155		
Qy 934	actaccctaatcaaaatagatccagcactgaagataaaaaacaaaagtgaactgca	993		
Db 156	ACTACCCCTAATCAAAATAGATCCAGCACTGAAGATAAAAACCAAAAAGTGAATCTGCA	215		
Qy 994	gaccaatgtgctaataagatgactaggaataaagaactccattcacttgcgaagccttt	1053		
Db 216	GACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTCTCTGCAAGGCTTTT	275		
Qy 1054	gttttgataagcaagaasaacaatgcctctggttcccttcaatagcatgtcaagtga	1113		
Db 276	GTTTITGATAAAGCAAGAAAACAATGCCTCTGGTTCCTTCAATAGCATGTCAAGTGGA	335		
Qy 1114	gtgaaaaagaatttggccatgaatttgacctctatgaaaaaagaactacattagaaac	1173		
Db 336	GTGAAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAAAAAAGAACTACATTAGAAAC	395		
Qy 1174	tgcattgtgaagagcagctacaagggaacagtatctatcactaagagtggcatc	1233		
Db 396	TGCAATCATGTGTAAAGACGACGCTACAAAGGAAACAGTATCTWTCTACTAAGAGTGGCATC	455		
Qy 1234	aaatgcagccctggagtctccatgataccacacagaacacagctatcgggggtaaagacct	1293		
Db 456	AAATGTCAGCCCTGGAGTTCCATGATACCAACACGAACACAGCTATCGGGGTAAAGAGCTA	515		
Qy 1294	caggaaaactactgtcgaaatcctcgaggggaagaaggggaccctggttttcacaagc	1353		
Db 516	CAGGAAAACTACTGTGAAATTCCTCGASGGGAAGAGGGGACCCTGTTGTTTCCAAAGC	575		
Qy 1354	aatccagaggtacgtctacgaagctgtgacattcctcagttctcagaagttgaaat	1408		
Db 576	AATCCAGAGGTACGCTACGAAGTCTGTGACATTCTCTCAGTCTTCAGAAAGTTAAT	630		

RESULT 5

BF383652

Query Match 25.3%; Score 428.4; DB 148; Length 652;
Best Local Similarity 81.3%; Pred. No. 1.7e-105;
Matches 509; Conservative 0; Mismatches 116; Indels 1;

Sequence	Matches	Mismatches	Indels	Gaps
Conservative	509	0	116	1
Indels	116	0	1	1

1 atgggggtgctcccaactcctctgctctctgactcaatgcttaggggtccctgggcaagcgc 60
28 ATGGGGTGGCTCGCAGCTTCCTGCTCTCTGGTACGTGTCAAGGCGCTCTGGCAGCGC 87
61 tcgcattgaatgacttccaaagtctccgggacagagactacagcacctgctacatgcg 120
88 TCACCATTGAATGACTTCCAGCTGTCCGGGGGCACAGAGTTAAGGAACCTGTTACACACA 147
121 tgggtgccggggccttggcagaggatgtggcagatgctgaagagtgtgctgctgcgtctg 180
148 GCGGTGCCGGGGCCATGCGCAGGAGGATGTGCAGATGCTGAGGAGTGTCTAGGCGCTGT 207
181 gggcccttaatgactgccgggcttccactacaacgctgagcagccatggttgcacaactg 240
208 GGGCCCCCTTCGTGGACTGTGGGGCTTCCACTACAACATGAGCAGCCATGGTTGCCAGCTG 267
241 ctgcattgactcaactgcgcacacgaggtcgcgcgttctctgggcgtctgtgacctc 300
268 CTGCGGTGGACCCAGCACCTCGCTGCACACACAGCTATACCACCTCGAGTCTGTGCCATCTC 327
301 ttccagaagaagactacgtacgagacctgcatacgaacatgggtgttgacctcggggc 360
328 TTCCACAGAAGAAGATTATGTGCGGACCTGCTATTATGACAAATGGGTCTAGCTACCGGGC 387
361 accatggccacgcgcgtgggtggcctcctcgcagcgttgagccacaaagtctccggaat 420
388 ACTGTGGCCAGCAGCAGCTGTGGCTCTGCCCTGCCAAGCCTGGAGTGCAGGTTCCCCAAT 447
421 gatcaaatgatacgcgccactctcgggaatggcctggaagaaactctgcogtaacct 480
448 GACCACAAGTATACGCCCCACGCCCAAAGAATGGCTGGAAGAGAACTTCTGTAGGAACCTC 507
481 gatggcaccocggaggtccttgggtgtatacaacaagaccctctgtgcgcttccagagc 540
508 GATGGGGATCCAGAGGTCCCTGGTGTGTACACAACAAACCGCAGTGTGCGTTTCCAGAGC 567
541 tgcgcatacaaatccttcggggagcgcgctgtctggtgcaatggcgaggaataccgc 600
568 TGTGGCATCAAAACCTCGCGGGAGGCTG-TGTGTCTGTGCACACGCTGAGGATTACCGT 626
601 ggcgcggttagaccgcacggagtcagg 626
627 GCGCAGGTAGACGTTACAGAGCTCAGG 652

RESULT 10
 .546856
 CCUCS
 DEFINITION AL546856 844 bp mRNA EST 16-FEB-2001
 AL546856 LTI_NFL006_PL2 Homo sapiens cDNA clone CSODI026VL03 5
 prime. mRNA sequence.
 CESSION AL546856
 ERSION AL546856.1 GI:12880379
 WORDS EST.
 URCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 844)
 LI,W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1 .844
 FEATURES
 source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YL03"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/notes="Vector: pcMVSPORT 6; Site.1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
301 a 152 c 57 g 304 t 30 others

```

	Query Match	24.5%	Score 414.2	DB 106	Length 844	
	Best Local Similarity	83.6%	Pred. No. 1.3e-101			
	Matches 460	Conservative	20	Mismatches 55	Indels 15	Gaps 3
Qy	874	tctctagaggacaaagaaaagaagaatacaattcatgaattcaaaaaatcagcaag	933			
Db	14	TATCAGAGGACAAAGAAAAGAAATACAATTCAATTCAGTAATTCAAAATACTAGCAAG	73			
Qy	934	actaccctaatacaaaatagatccagcactgaagataaaacccaaaaagtgaatactgca	993			
Db	74	ACTACCCTAATCAAAATAGATCCAGCACTGAAGATAAAAAACCAAAAAAGTCAATACTGCA	133			
Qy	994	gaccaatgtgctaatagatgtactaggaaataaaaggacttccattcacttgcaaggctttt	1053			
Db	134	GACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTACCTTCGAAGGCTTTT	193			
Qy	1054	gttttataagcaagaaacaatgcctctgtgtcccttcaatagcatgtcaagtga	1113			
Db	194	GTTTTTGCATAAGCAAGAAACAATGCCTCTGGTTCCTTCCTCAATAGCATGCTCAAGTGA	253			
Qy	1114	gtaaaaaagaattggcccatgaattgacctctatgaaaaaagaagcactcattagaaac	1173			
Db	254	GTKAAAAAGAAATTTKKCCATGAATTTKACCTCTATKAAAAAAGAGACTACATTAGAAAC	313			
Qy	1174	tgcattattgtaagacagcagctacaagggacagtatctatcactaagagtggcac	1233			
Db	314	TGCATCATTTKGTAAAGGACKSAKSTACATTTTASATTTATCTATCASAATATTTTTCATC	373			
Qy	1234	aaatgtcagccctggagtctcatgatataccacacgaacac-----agctat	1278			
Db	374	AAATTTCATCCCTTTAATTTCCATTATACCACACTAAACACTCTTTTTTCSTTCTATCTAT	433			
Qy	1279	cggggtaaaagacctacagggaaaactactgtcgaaatcctcgaggggaagaaggggacc	1338			
Db	434	CTTKTTAAATACCTTACATTAATAAGTACTTTTCTAAATCCTCTABTTTAAATMTTTTACCC	493			
Qy	1339	tgggtttcacaagcaatccagagggtacgcgtacgaagctcgtgacactccctcagtgcca	1398			
Db	494	TTTTTTTTTCAACATCAATCCATATTTACTCTACTCAATNCTTTTASATSSSATTTTISA	553			
Qy	1399	gaagttgaat	1408			
Db	554	TAAATTTAAAT	563			

RESULT	11
BE723534	
LOCUS	BE723534 543 bp mRNA EST
DEFINITION	193407 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION	BE723534
VERSION	BE723534.1 GI:10124839
KEYWORDS	EST.
SOURCE	COW.
ORGANISM	Bos taurus

Db 497 AGGGAAGAGGGGACCCCTGGTGTTCACAAAGCAATCCAGAGGTACGCTACGAAGTCTG 438

QY 1380 tgacattcttcgaagtttcgaagttgaatgcacatgacccgcgaatgggagagttatcgagg 1439

Db 437 TGCAATTCCTCAGTGTTCAGAAGTNGAATGCATGACCTGCAATGGGAGAGTTCGAGG 378

QY 1440 ttcattgatcatcagaatcagcaagattgtcagcgtggatcatcagacacaca 1499

Db 377 TCATCAT-GATCATACAGAATCAGCAAGA-TTGTACGCGT-GGATCATCAGACACCACA 321

QY 1500 ccggcacaattcttgcctgaagatcccgacaaaggttttgcattatattattccg 1559

Db 320 CCGGCAAAATTCCTGCTGAAGATATCCGACAAGGCTTTCATGATTAATTTCGCG 261

QY 1560 caatcccgatggcagcgagggcgcattgtctatactcttgacccctcacaccgcctggga 1619

Db 260 CAATCCCGATGGCAGCGGCGCATGGTGTATACCTTTGACCCCTCACACCCGCTGGGA 201

QY 1620 gtactgtgcaattaaaaacatgcg 1642

Db 200 GTACTGTGCAATTAANACATGCG 178

RESULT 13

AA037738 506 bp mRNA EST 01-FEB-1997

LOCUS zf03b09.r1 Soares_fetal_heart_NDH19W Homo sapiens cDNA clone

DEFINITION IMAGE:375833 5' similar to gb:M73239 HEPATOCYTE GROWTH FACTOR

PRECURSOR (HUMAN), mRNA sequence.

ACCESSION AA037738

VERSION AA037738.1 GI:1512912

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Hillier, L., Lennon, G., Becker, M., Donald, M.F., Chiapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M.

Generation and analysis of 280,000 human expressed sequence tags

Genome Res. 6 (9), 807-828 (1996)

97044478

CONTACT: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1135 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 496.

TITLE
JOURNAL
MEDLINE
COMMENTFEATURES
source

1. .506

/organism="Homo sapiens"

/db_xref="GB:1284089"

/db_xref="taxon:9606"

/clone="IMAGE:375833"

/clone_lib="Soares_fetal_heart_NDH19W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

/notes="Organ: heart; Vector: pT7m3D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of a modified pT7m3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung Nbh19W."

BASE COUNT 172 a 124 c 98 g 112 t

ORIGIN

Query Match 20.7%; Score 350.2; DB 1; Length 506;

Best Local Similarity 98.9%; Pred. No. 2.9e-84;

Matches 363; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 874 tctctagagggacaaagaaaagaaatacaattcatgaattcaaaaataatcagcaaaag 933

Db 140 TATGCAGAGGGACAAAGGAAAGAAATACAATTTCATGAATTCAAAAATCAGCAAG 199

QY 934 actacccttaataatagatccagcactgaagataaaacccaaaagtgataactgca 993

Db 200 ACTACCCTAATCAAAATAGATCCAGCACTGAAGATAAAAACCAAAAGTGAATACTGCA 259

QY 994 gaccaatgtcctaagatgtactaggaataaaggacttccattccattgcaaggctttt 1053

Db 260 GACCAATGTCTAATAGATGTACTAGGAATAAAGGACTTCCATTCTACTTGCAGGCTTT 319

QY 1054 gttttgataagcagaagaaacaatgcctctgttcccttcaatagcatgtcaagtggga 1113

Db 320 GTTTTGTATAAGCAAGAAACAATGCCCTCTGCTCCCTTCAATAGCATGTCAAGTGA 379

QY 1114 gtgaaaaagaatttggcccatgaatttgacctctatgaaacaaaagactacattagaaac 1173

Db 380 GTGAAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAAAAACAAAGACTACATTAGAAC 439

QY 1174 t-gcatcttgtaagacgcagctcacagggaacagtagtatctatcactaagagtggtcat 1232

Db 440 TGGCATCATTTGTTAAAGGAGCGCAGCTACAGGGAACAGTATCTATCTACTAAGAGTGGCAT 499

QY 1233 caaatgt 1239

Db 500 CAAATGT 506

RESULT 14

BE755132 518 bp mRNA EST 15-SEP-2000

LOCUS BE755132

DEFINITION 208667 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE755132

VERSION BE755132.1 GI:10169124

KEYWORDS EST.

SOURCE COW.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 518)

AUTHORS Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Laegreid, W.W. and Keeler, J.W.

Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle

Unpublished (2000)

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:21:43 ; Search time 4515.12 Seconds
(without alignments)
5909.448 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtgggtgacaaactoct.....accaccaccaccaccactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pll:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_rod:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vl1:*
59: gb_vl2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_rod:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1725	100.0	1725	9	AX019527 Sequence
2	996.6	57.8	1692	9	AX019529 Sequence
3	857.8	49.7	2190	9	AX019515 Sequence
4	857.8	49.7	2190	9	AX019517 Sequence
5	857.8	49.7	2199	10	E05008 HLC2 cDNA s
6	857.8	49.7	2199	10	E16560 Human HLC3
7	857.8	49.7	2289	10	E12582 CDNA encodi
8	857.8	49.7	2704	97	M73240 Human (clon

```
Qy 1141 agtgagtgaaaaaagatttggccatgaatttgacctcttatgaaacaaagactacatt 1200
|||||
Db 1141 AGTGGAGTGAaaaaaagatttggccatgaatttgcctcttatgaaacaaagactacatt 1200

Qy 1201 agaaactgcatttgtaaaagcgcagctacaaaggaacagtatctatcaactaagagt 1260
|||||
Db 1201 AGAACTGCATTCATTTGTTAAAGGAGCGAGCTACAAAGGAACAGTATCTATCACAAGAGT 1260

Qy 1261 ggcataaattcagccctggagttccatgataccacacacacagctatcggggtaaa 1320
|||||
Db 1261 GGCATCAAAATGTCAGCCCTGGAGTTCCATGATACACACGACACAGCTATCGGGGTAAA 1320

Qy 1321 gacctacaggaactactgcgaattcctcagggagggggaaggggacccctggtgttc 1380
|||||
Db 1321 GACCTACAGGAATACTACTGTCGAAATCCTCGAGGGGAAGAGGGGACCCCTGGTGTTC 1380

Qy 1381 acaagcaatccagaggtaacctcagaagctgtgacattcctcagtttcaagaagttaa 1440
|||||
Db 1381 ACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAGTGTTCAGAAGTTGAA 1440

Qy 1441 tgcatacctcaatgggagagttatcgaggtctctggtatcatatcagaaatcaggaag 1500
|||||
Db 1441 TGCATGACCTCAATGGGAGAGTTATCGAGTCTCATGTGATCATACAGATTCAGGCAAG 1500

Qy 1501 atttgtcagcgtgggatcatcagacaccacacccggcacaattcttgcctgaagatat 1560
|||||
Db 1501 ATTGTGACGCTGGGATCATCAGACACACACACCCGGCACAAATCTTTCCTGAAAGATAT 1560

Qy 1561 cccgacaagggcttgatgataattattgcgcgaatcccgatggccagcagggccatgg 1620
|||||
Db 1561 CCCGACAAGGGCTTTGATGATAATTATTCGCGCAATCCGATGGCCAGCCAGGCCCATGG 1620

Qy 1621 tgcataccttgacctccacacccgtggagtgactgtgcaattaaacacatgcgtgac 1680
|||||
Db 1621 TGCATACTCTTGACCTCAGCCCTCAGCCGCTGGGAGTACTGTGCAATTAACATATGCGCTGAC 1680

Qy 1681 aaagtgacgacgacgacacacacacacacacacacacacacacacacacacacacac 1725
|||||
Db 1681 AAAGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1725

RESULT 2
AX019529
LOCUS AX019529 1692 bp DNA PAT 07-SEP-2000
DEFINITION Sequence 3 from Patent WO9938967.
ACCESSION AX019529
VERSION AX019529.1 GI:10043446
KEYWORDS
SOURCE synthetic construct.
ORGANISM
REFERENCE 1 (bases 1 to 1692)
AUTHORS Medico,E., Collesi,C., Comoglio,P., Michieli,P. and Caselli,G.
TITLE Recombinant proteins derived from hgf and msp
JOURNAL Patent: WO 9938967-A 3 05-AUG-1999;
DOMPE SPA (IT); MEDICO ENZO (IT); COLLESI CHIARA (IT); COMOGLIO
PAOLO (IT); MICHELIEL PAOLO (IT); CASELLI GIANFRANCO (IT)
FEATURES
source
1..1692
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="RECOMBINANT"
BASE COUNT 442 a 457 c 459 g 334 t
ORIGIN
Query Match 57.8%; Score 996.6; DB 9; Length 1692;
Best Local Similarity 83.2%; Pred. No. 1.6e-255;
Matches 1166; Conservative 0; Mismatches 214; Indels 21; Gaps 2;
Qy 346 ttgacctctatgaaacaaagactacattagaactgcatttgtaaaagcagcagc 405
|||||
```

Db 61 CTGCTCCCATCGCCATCCCTATGAGAGGACAAAGGAAAGAGAAATACAAATTCAT 120
QY 121 gaattcaaaaaatcagcaaaagactacccttaatacaaaatagatccagcactgaagataaaa 180
Db 121 GAATTCAAAAAATCAGCAAAAGACTACCCCTAATCAAAATAGATCCAGCAGCTGAAGATAAAA 180
QY 181 accaaaaagtgaatactgacagaccactgtgtaataagatgtactagggaataaaaggactt 240
Db 181 ACCAAAAAGTGAATACGTGACAGACCAATGTCTAATAGATGTACTAGGAATAAAGGACTT 240
QY 241 ccattcactgcaaggctttttttttgataaagcaagaacaatgcctctggttcccc 300
Db 241 CCATTCACCTGCAAGCCTTTTGTGATAAGCAAGAAACAATGCTCTGGTTCCTCC 300
QY 301 ttaatatagcatgcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 301 TTCAATAGCATCTCAAGTGGAGTGAAAGAAAGAAATTTGGCCATGAATTTGACCTCTATGAA 360
QY 361 acaaaagactacattagaaactgcatacttgtaaaaggcagcagctacaaaggaacagta 420
Db 361 AAAAAAGACTACATTAGAACTGCATCAATTTGTAAGGAGCGAGCTACAAAGGAACAGTA 420
QY 421 tctatcactaaagtgagcatcaaatgtcagccctgaggttccataaccacacaaacac 480
Db 421 TCTATCCTTAGAGTGGCATCAAAATGTCAGCCCTGGAGTTCATGATACACACGAAACAC 480
QY 481 agctatcggggttaagacacacacagaaactactgtcgaatcctcaggggaaagaagg 540
Db 481 AGCTATCGGGGTAAAGACCTACAGGAAAACTACTGTGCAAAATCCTCGAGGGGAAGAGG 540
QY 541 ggaacctgggtttcacaaagaatccagaggtacgtacagagctgtgacattctctcag 600
Db 541 GGACCTGGTGTTCACAAAGCAATCCAGAGGTACGCTACGAAAGTCTGTGACATTCCTCAG 600
QY 601 tgttcagaagtgaatgatgacctgcaatgggagagttatcgaggtctcagatgatcat 660
Db 601 TGTTTCAAGTTGAATGATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCAT 660
QY 661 acagaatcagggaagatttgcagcgtgggatcatcagacacacacacgggcacaaattc 720
Db 661 ACAGAATCAGGCAAGATTGTGAGCGCTGGGATCATCAGACACACACCGGCACAAATTC 720
QY 721 ttgctgtaaaagatacccgacaaggcttgatgataattatgcgcaatccccgatgac 780
Db 721 TTGCTGTAAAGATATCCCGACAGGGCTTTGATGATAATTTATTTGGCGCAATCCCGATGGC 780
QY 781 cagccgagccatgctctactacttgaacctcaccctcagccgctggagtagtgcatt 840
Db 781 CAGCCGAGCCCATGGTGTATCTCTTGACCTTCACCCCGCTGGGAGTACTGTGCAATT 840
QY 841 aaaaacatgcgctgacaaaagct 861
Db 841 AAAACATGCGCTGACAATACT 861
RESULT 5
LOCUS E05008 2199 bp RNA PAT 29-SEP-1997
DEFINITION HLC2 cDNA sequence coding for human hepatoparenchymal cell growth factor(HGF).
ACCESSION E05008
VERSION E05008.1 GI:2173203
KEYWORDS JP 1993111383-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2199)
AUTHORS Nakamura,T., Hagiya,M., Seki,T., Shimonishi,M., Shimizu,S., Inohara,I., Sakaguchi,M. and Asami,O.
TITLE RECOMBINANT HUMAN HEPATOCYTE GROWTH FACTOR AND ITS PRODUCTION
JOURNAL Patent: JP 1993111383-A 2 07-MAY-1993;

NAKAMURA TOSHIICHI
OS Homo sapiens (human)
PN JP 1993111383-A/2
PD 07-MAY-1993
PF 06-JUN-1991 JP 1991163485
PR 11-JUN-1990 JP 90P 152474
PI NAKAMURA TOSHIICHI, HAGIYA MICHIO, SEKI TATSUYA, PI SHIMONISHI MANABU,
PI SHIMIZU SHIN, INOHARA IZUMI, SAKAGUCHI MARIKO, ASAMI OSAMU PC C12N15/16.C07K13/00.C12N5/10.C12N15/85.C12P21/02//A61K37/02, PC A61K37/24.
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC *source: tissue_type=Blood;
CC *source: cell_type=Leukocyte;
PH Key Location/Qualifiers
FT 5'UTR 1..17
FT sig_peptide 18..110
FT 111..2172 /product='human HGF signal peptide' FT CDS
FT /product='human HGF'.
FEATURES
source
Location/Qualifiers
1..2199 /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 678 a 464 c 505 g 552 t
ORIGIN
Query Match 49.7%; Score 857.8; DB 10; Length 2199;
Best Local Similarity 99.8%; Pred. No. 2e-218;
Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 atgtggttaacaaactctgcagccctgctgctgagagggacaaaggaagaaatacaattcat 120
Db 28 ATGTGGTGACCAAACTCCCTGCCAGCCCTGCTGCTGAGCAATGCTCCTCCTCATCTCCTC 87
QY 61 ctgtctcccatcgccatccctctatgcagagggacaaaggaagaaatacaattcat 120
Db 88 CTGCTCCCATCGCCATCCCTATGCAAGAGGACAAAGGAAAGAAATACAAATTCAT 147
QY 121 gaattcaaaaaatcagcaaaagactacccttaatacaaaatagatccagcactgaagataaaa 180
Db 148 GAATTCAAAAAATCAGCAAAAGACTACCCCTAATCAAAATAGATCCAGCAGCTGAAGATAAA 207
QY 181 accaaaaagtgaatactgacagaccactgtgctaaatagatgtactaggaaataaaggactt 240
Db 208 ACCAAAAAGTGAATCTGAGACCAATGCTAATAGATGTACTAGGAATAAAGGACTT 267
QY 241 ccattcactgcaaggctttttttgataaagcaagaacaatgcctctggttcccc 300
Db 268 CCATTCTACTGCAAGGCTTTTGTGATAAAGCAAGAAACAATGCCTCTGTTCCCTC 327
QY 301 ttaatatagcatgcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 328 TTCAATAGCATCTCAAGTGGAGTGAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAA 387
QY 361 acaaaagactacattagaaactgcatactggtaaaagagcagcagctacaaaggaacagta 420
Db 388 ACAAAGACTACATTAGAAACTGCATCATTGGTGAAGGACCGCAGCTACAAAGGAACAGTA 447
QY 421 tctatcactaaagtgagcatcaaatgtcagccctggagttccatgataccacacgaacac 480
Db 448 TCTATCCTTAAGAGTGGCATCAAAATGTCAGCCCTGGAGTTCATGATACACGAAACAC 507
QY 481 agctatcggggttaagacacacacagaaactactgtcgaatcctcaggggaaagaagg 540
Db 508 AGCTATCGGGGTAAAGACCTACAGGAAACTACTGTGCAAAATCCTCGAGGGGAAGAGGG 567
QY 541 ggaacctgggtttcacaaagaatccagaggtacgcgtacgaagtctgtgacattctcag 600
|||||

CONSISTING OF THE SAME AND PHARMACEUTICAL PREPARATION WITH THE SAME

AS ACTIVE INGREDIENT

Patent: JP 1997025297-A 1 28-JAN-1997;

SNOW BRAND MILK PROD CO LTD

OS Homo sapiens (human)

PN JP 1997025297-A/1

PD 28-JAN-1997

PF 09-MAR-1990 JP 1996191458

PI TOO KANJI, MITSUDA SHINJIRO, SHIMA NOBUYUKI, ITAGAKI KOJI, PI

NAGAO MASAYA

PC C07K14/52.A61K38/00.A61K38/00.A61K38/00.A61K38/00.A61K38/00,

PC A61K38/00,

PC A61K38/00.C12N15/09//C12P21/02.(C12P21/02,C12R1.91); CC

strandedness: Double;

CC topology: linear;

CC hypothetical: No;

CC anti-sense: No;

FH Key Location/Qualifiers

FT source

1..2289

/organism='Homo sapiens'

/tissue_type='Lung'

/cell_type='Fibroblast'

/cell_line='IMR-90'

FT 5'UTR

FT CDS

78..2249

/product='glycoprotein,TCF-2' FT sig_peptide

FT 78..918

/product='signal peptide of

glycoprotein,TCF-2'

FT mat_peptide

919..2246

/product='glycoprotein,TCF-2' FT 3'UTR

FT 2250..2289.

Location/Qualifiers

source

1..2289

/organism='Homo sapiens'

/db_xref='taxon:9606'

BASE COUNT 703 a 491 c 523 g 572 t

ORIGIN

Query Match 49.78; Score 857.8; DB 10; Length 2289;

Best Local Similarity 99.88; Pred. No. 2e-218;

Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 atgtgggtgacaaactctgccagccctgctgtcgcagcatgtcctctcctcatctcctc 60

Db 78 ATGTGGGTGACCAACTCTGCCAGCCCTGCTGCTGCGAGCATGCTCTCTCATCTCCTC 137

Qy 61 ctgctcccatcgccatccccctgtcagaggagacaaagaaagaaatacaattcat 120

Db 138 CTGCTCCCATCGCATCCCTATGCGAGGSGACAAAGAAAGAAATACRAATTCAT 197

Qy 121 gaattcaaaaaatcagaaagactaccctaatcaaaatagatcagacactagaataaaa 180

Db 198 GAATTCAAAAAATCAGCAAGACTACCCCTAATCAAAATAGATCAGCACTGAAGATAAA 257

Qy 181 accaaaaagtgaactcagacccaatgtctaataagatactagaaataaaggactt 240

Db 258 ACCAAAAAGTGAATCTGCGAGACCAATGCTATATAGTACTAGGAATAAAGGACTT 317

Qy 241 ccattcacttcaaggctttttttttgataaaagcaaaacaatgctctgtgtccccc 300

Db 318 CCATTCACCTCAAGGCTTTTGTGTTTGTGATAAGCAAGAAACAATGCTCTGGTTCCCC 377

Qy 301 ttcaatagcatgtcaagtggagtgaaagaaagaaatttgcccatgaattgacctctatgaa 360

Db 378 TTCAATAGCATGTCAAGTGGAGTGAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAA 437

Qy 361 aacaaagactacattagaacctcatcattgtgtaagagcagcagctacaaggaacagta 420

Db 438 ANCAAAGACTACATTAGAAACTGTCATCTATTGGTAAGGACGACGCTACAAGGGAACAGTA 497

Qy 421 tctatcactaaagtgagcatcaaaatgtcagccctgaggtccatgatataccacacgaac 480
Db 498 TCTATCACTAAGAGTGGCATCAAAATGTGAGCCTGTGAGTTCCATGATACACAGAACAC 557
Qy 481 agctatcggggttaaagaccctacaggaataactactgtcgaatactcctcgaggggaagagg 540
Db 558 AGCTATCGGGGTAAAGACCTACAGGAAATACTACTGTGAAATCCTCGAGGGGAAGAGG 617
Qy 541 ggaacctggtgtttcacagaatccagagggtacgtacagaaagctctgtgacattctctcag 600
Db 618 GGACCCCTGCTGTTCACAAAGCAATCCAGAGGTACGCTAGCAAGTCTGTGACATTCCTCAG 677
Qy 601 ttttcagaagtgtgaatgacatgacctgcaatggagagattatcagaggtctcatggatcat 660
Db 678 TGTTCAGAAAGTTGAATGTCATGACCTGCAATGGGAGAGTTATCGAGGTCCTCATGGATCAT 737
Qy 661 acagaatcagggcaagattttgtcagcgtgggacatcagacaccacacccgagcaaatc 720
Db 738 ACAGAATCAGGCAAGATTTGTGACGCTGGGATCATCAGACACCACCGGCACAATTC 797
Qy 721 ttgctgaagatatcccgacaaaggcttttgatgataataattgcccgaatcccgatggc 780
Db 798 TTGCTGAAAGATATCCCGACAAGGCTTTGATGATAAATATTGCCGAATCCCGATGGC 857
Qy 781 cagccgagggccatggtgtctatactcttgacctcacaccgcctggagagtagtctgaatt 840
Db 858 CAGCCGAGGCCATGCTGTATATCTTTGACCTCAGACCCGCTGGAGTAGTACTGTGAATT 917
Qy 841 aaacatcgctgcagcaaaagct 861
Db 918 AAACATGCGCTGACAATACT 938

RESULT 8

HUMSCFA2

LOCUS

DEFINITION

HUMAN (clone SF2) hepatocyte growth factor (HGF) mRNA, complete

cds.

ACCESSION

M73240

VERSION

M73240.1 GI:727162

KEYWORDS

hepatocyte growth factor; scatter factor.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2704)

AUTHORS

Weidner,K.M., Arakaki,N., Hartmann,G., Vandekerckhove,J.S.,, Hishida,T., Weingart,S., Kieder,H., Fonatsch,C., Tsubouchi,H., Hishida,T., Daikuhara,Y. and Birchmeier,W.

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE

Evidence for the identity of human scatter factor and human hepatocyte growth factor

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 88 (16), 7001-7005 (1991)

MEDLINE

91334393

COMMENT

On Mar 24, 1995 this sequence version replaced gi:337937.

FEATURES

Location/Qualifiers

1..2704

/organism='Homo sapiens'

/db_xref='taxon:9606'

/clone='SF2'

/cell_line='MRC-5'

/dev_stage='fibroblast'

/map='7q21'

<1..2689

/gene='HGF'

/note='G00-127-524'

1..2689

/gene='HGF'

1..65

/gene='HGF'

/note='G00-127-524'

66..2237

/gene='HGF'

CDS

QY 241 ccattcactgcaaggcttittgttttttataaagcaaaacaaatgcctctgttcccc 300
|||||
Db 241 CCATTCTACTGCAAGGCTTTTGTGTAAGACAAAGAAACAAATGCTCTGTGTTCCCC 300
QY 301 ttcaatagcatgcaagtgaagtgaaagaaatattggcccatgaatttgacctctatgaa 360
|||||
Db 301 TTCAATAGCATGCAAGTGGAGTGAAAGAAAGAAATTTGGCCATGAATTTGACCTCTATGAA 360
QY 361 acaaaagactacattagaaactgcatttggttaagagacgcagctacaaaggaacagta 420
|||||
Db 361 AACAAAGACTACATTTAGAAACTGCAATGCTGTAAGAGACGAGCTACAAGGAACAGTA 420
QY 421 tctatactaagagtgcatcaaatgctcagccctggagtttccatgatataccacagaacac 480
|||||
Db 421 TCTATCTACTAAGAGTGCATCAAAATGTCAGCCCTGGAGTTCCATGATACCAACGAAACAC 480
QY 481 agctatcggggttaagacccctacagaaactactgtcgaaatcctcgagggggaagagg 540
|||||
Db 481 AGCTATCGGGTAAAGACCTACAGAAACTACTGTGCGAAATATCTCGAGGGGAAGAGGG 540
QY 541 ggaacctgggtttccaaagcaatccagaggtacgctacgaagtctgtgacatctcctcag 600
|||||
Db 541 GGACCCCTGGTGTTCACAAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAG 600
QY 601 tgttcagaagtgaatgcagac 660
|||||
Db 601 TGTTCAGAAGTTGAATGCATGACCTGCAATGGGAGAGGTTTATCGAGGTCTCATGATCAT 660
QY 661 acagaatcaggcaagatttgcagcgcctgggatcatcagacacacacacacacacacacac 720
|||||
Db 661 ACAGAATCAGGCAAGATTGTGCAGCGTGGGATCATCATGACACACACACACACCAAAATTC 720
QY 721 ttgacctgaagatataccgacaaagggtttgatgataattattgccgcaatccccgatgac 780
|||||
Db 721 TTGCTTGAAGATATCCCGAAGAGGCTTTGATGATAAATTTGCCGCAATCCCGATGGC 780
QY 781 cagccgaggccatgctatactactgtgacccctcacaccccgctgggagtaactgtgcaatt 840
|||||
Db 781 CAGCCGAGGCCATGGTCTACTTGTGACCCCTCACACCCCGCTGGAGTACTGTGCAATT 840
QY 841 aaacatgcgtgacaaagct 861
|||||
Db 841 AAAACATGCGCTGACAACTACT 861

RESULT 10
I55856 155856 2288 bp DNA PAT 07-OCT-1997
LOCUS Sequence 4 from patent US 5648233.
DEFINITION I55856
ACCESSION I55856
VERSION I55856.1 GI:2476650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2288)
AUTHORS Yamaguchi,K., Shima,N., Murakami,A., Goto,M., Tsuda,E.,
Masunaga,H., Takahira,R., Ogaki,F., Ueda,M. and Higashio,K.
TITLE Modified tumor cytotoxic factor (TCF) and DNA encoding such
JOURNAL Patent: US 5648233-A 4 15-JUL-1997;
FEATURES Location/Qualifiers
source
1. .2288
/organism="unknown"
BASE COUNT 702 a 491 c 524 g 571 t
ORIGIN

Query Match 49.0%; Score 845.8; DB 10; Length 2288;
Best Local Similarity 99.7%; Pred. No. 3.3e-215;
Matches 858; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 atgtgggtgacaaactctgcagccctgctgctgagcagatgctcctcctgcatcctc 60
|||||

Db 78 ATGTGGTGACCAAACTCCTGCCAGCCCTGCTGCTGCAGCATGCTCCTCTCGATCTCCTC 137
QY 61 ctgctcccatcgccatccccctatgcagagggaacaaagaaagaaatacaaatcat 120
|||||
Db 138 CTGCTCCCATCCCATCCCTATGCGAGGGACAAAGGAAAGAAATACAAATTCAT 197
QY 121 gaattcaaaaaatcagcaaaagactacccttaatacaaaaatagatccagcactgaagataaa 180
|||||
Db 198 GAATTTCAAAAAATCAGCAAAAGACTACCCCTTAATCAAAA-AGATCCAGCAGCTGAAGATAAAA 256
QY 181 accaaaaagtgaatactgcagacaccaatgtgtaataagatgtactaggataaagagactt 240
|||||
Db 257 ACCAAAAAAGTGAATGTCAGACACCAATGTGTAATAGATGTAAGATAAAGGACTT 316
QY 241 ccattcacttgcaaggcttttttttgataaagcaaaacaaatgcctctggttcccc 300
|||||
Db 317 CCATTCTACTGCAAGGCTTTTGTGTTGATAAAGCAAGAAACAAATGCTCTGTTCCCC 376
QY 301 ttcaatagcatgcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
|||||
Db 377 TTCAATAGCATGCTCAAGTGGAGTGAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAA 436
QY 361 acaaaagactacattagaaactgcattgttaagagacgcagctacaaaggaacagta 420
|||||
Db 437 AACAAAGACTACATTAGAAACTGCAATGTTGGTAAAGGACGACGTACAAAGGAACAGTA 496
QY 421 tctatactaagtgatgcaatcaaatgctcagccctggagtgtccatgataccacacgaacac 480
|||||
Db 497 TCTATCTACTAAGTGGCATCAAAATGTCAGCCCTGGAGTTCCATGATACCAACGAAACAC 556
QY 481 agctatcggggttaagacccctacagaaactactgtcgaaatcctcgagggggaagagg 540
|||||
Db 557 AGCTATCGGGGTAAAGACCTACAGGAAAACTACTGTGCAAAATCCTCGAGGGGAAGAGGG 616
QY 541 ggaacctgggtttccaaagcaatccagaggtacgctacgaagtctgtgacatctcctcag 600
|||||
Db 617 GGACCCCTGGTGTTCACAAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAG 676
QY 601 tttcagaagtgtgaatgcagcctgcaatggggagagttatcgaggtctcatggtatcat 660
|||||
Db 677 TGTTCAGAAGTTGAATGTCATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCAT 736
QY 661 acagaatcaggcaagatttgcagcgcctgggatcatcagacacacacacacacacacacac 720
|||||
Db 737 ACAGAATCAGGCAAGATTGTGAGCGCTGGGATCATCAGACACACACCGGCAAAATTC 796
QY 721 ttgcttgaagatataccgacaaaggctttgatgataattattgccgcaatccccgatgac 780
|||||
Db 797 TTGCTTGAAGATATCCCGAAGGCTTTGATGATAAATTTGCCGCAATCCCGATGGC 856
QY 781 cagccgaggccatggtgctatactcttgacctcacaccccgctgggagtaactgtgcaatt 840
|||||
Db 857 CAGCCGAGGCCATGGTGTATATCTCTTGACCCCTCACACCCCGCTGGAGTACTGTGCAATT 916
QY 841 aaacatgcgtgacaaagct 861
|||||
Db 917 AAAACATGCGCTGACAACTACT 937

RESULT 11
E06576 E06576 2184 bp RNA PAT 29-SEP-1997
LOCUS
DEFINITION cDNA encoding human growth factor for endothelial cell.
ACCESSION E06576
VERSION E06576.1 GI:2174763
KEYWORDS JP 1994009691-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2184)
AUTHORS Shudo,T., Harada,K., Hirahara,I. and Adachi,M.
TITLE PROTEIN ORIGINATED FROM HUMAN

```

Db 61 CTGCTCCCATCGCCATCCCTATGAGAGGACAAAGAAAGAAATACAATTCAT 120
Qy 121 gaattcaaaaaatcagcaaaactaccctaatacaaaatagatccagcactgaagataaaa 180
Db 121 GAATTCAAAAAATCAGCAAAAGACTACCTAATCAAAATAGATCAGCAGCTGAAGATAAAA 180
Qy 181 accaaaaaagtgaatactgcagacaaatgtctaataagatgtactaggaaataaaggactt 240
Db 181 ACCAAAAAAGTGAATCTGCAAGACCAATGTCTAATAGATCTACTAGGAATAAAGGACTT 240
Qy 241 ccattcactgcaaggctttgttttttgataaagcaagaaacaatgcctctggtccccc 300
Db 241 CCATTCACCTTCAGAGCTTTTGTGATAGCAAGAAACAATAGCTCTGTGTTCCCC 300
Qy 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgactcctctatgaa 360
Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAGCAATTTGGCCCATGAATTTGACCTCTATGAA 360
Qy 361 acaaaagactacattagaaactgcatcattgttgtaagagcagctacaagggaaacagta 420
Db 361 AACAAAGACTACATATAGAAACTGCATCATTTGTAAGAGCAGCAGCTACAAGGGAACAGTA 420
Qy 421 tctatcactaagatggcatcaaaatgtcagccctgaggttccatgataccacacagaacac 480
Db 421 TCTATCACTAAGAGTGGCATCAAAATGTCAGCCCTGGAGTTCATGATACCAACAGAACAC 480
Qy 481 -----agctatcggggttaagacacctacagggaaactactgtcgaatcct 525
Db 481 AGCTTTTTCCTTCAGCTATCGGGTAAAGACCTACAGGAAACTACTGTGCGAAATCCT 540
Qy 526 cgaagggaagaagggaagccctgtgtttcaagcaatccagagggtcgcgtacgaagtc 585
Db 541 CGAGGGGAAGAAGGGGACCCTGGTGTTCACAAAGCAATCCAGAGGTACGCTACGAAGTC 600
Qy 586 tgtacattcctcagttgtcagaagttgaatgcattgcatgacctgaatgggagagttatcga 645
Db 601 TGTGACATTCCTCAGTGTTCAGAGTGAATGCAATGCAATGGGAGAGTTATCGA 660
Qy 646 ggtctcatggatcatacagaaatcaggaagatttgcagcgtggatcatcagacacca 705
Db 661 GGTCTCATGGATCATACAGAAATCAGGCAAGATTTGTACGCGCTGGGATCATCAGACACCA 720
Qy 706 caccggcaaaaattctgctcgtgaagatatcccgacaaggccttggatgataattatgc 765
Db 721 CACCGGCACAAATTTTGCCTGAAAGATATCCGACACAGGGCTTTGATGATAATTTATTGC 780
Qy 766 cgaatccgatgcccagccagccatggtgctatactcttgaccttcacaccogctgg 825
Db 781 CGCAATCCCGATGCGCCAGCCGAGGCCATGGTGCATATCTCTTGACCCCTCACACCCGCTGG 840
Qy 826 gagtactgtgcaattaaacacatgcctgacaaaagct 861
Db 841 GAGTACTGTGCAATTAACACATGCGTGCACAACTACT 876

```

RESULT 13

E09626

ID E09626 standard; RNA; HUM; 2184 BP.

XX AC E09626;

SV E09626.1

XX XX

DT 07-OCT-1997 (Rel. 52, Created)

DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)

XX XX

DE DNA encoding human HGF.

XX XX

KW JP 1995179356-A/1.

XX XX

OS Homo sapiens (human)

XX XX

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.

XX [1]

RN 1-2184

RP Nakamura T., Matsumoto K.;

RT "EPITHELIOCYTE GROWTH PROMOTER";

RL Patent number JP1995179356-A/1, 18-JUL-1995.

RL NAKAMURA TOSHIICHI.

XX XX

OS Homo sapiens (human)

CC PN JP 1995179356-A/1

CC PD 18-JUL-1995

CC PF 28-DEC-1990 JP 1990419158

CC PI NAKAMURA TOSHIICHI, MATSUMOTO KUNIO

CC PC A61K38/00,A61K38/00//C12N15/16,C12P21/02,(C12P21/02,C12R1:91);

CC CC strandedness: Double;

CC CC topology: Linear;

CC CC Key

CC FH Location/Qualifiers

CC FH source

CC FT 1. 2184

CC FT /organism="Homo sapiens"

CC FT /tissue_type="liver"

CC FT mat_peptide

CC FT 163..1482

CC FT /product="HGF alpha chain"

CC FT mat_peptide

CC FT 1483..2184

CC FT /product="HGF beta chain"

XX Key

FH Location/Qualifiers

FH FT source

FH FT 1. 2184

FH FT /db_xref="taxon:9606"

FH FT /organism="Homo sapiens"

XX XX

SQ Sequence 2184 BP; 673 A; 455 C; 500 G; 556 T; 0 other;

Query Match

Best Local Similarity 48.3%; Score 832.8; DB 45; Length 2184;

Matches 859; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 atgtgggtgacaaactcctgcagccctgctgctgcagcatgtcctcctgcactctc 60

Db 1 ATGTGGGTGACCAAACTCCTCTGCGAGCCCTGCTGCTGCAGCATGTCTCTCTCTCCTC 60

Qy 61 ctgctccccatcgccatccccctatgcagagggaacaaaggaaagaaatacaaatcat 120

Db 61 CTGCTCCCATCGCCATCCCTATGCAGAGGACAAAGGAAAGAAATACAATTCAT 120

Qy 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

Db 121 GAATTCAAAAAATCAGCAAAAGACTACCCCTAATCAAAATAGATCCAGCATGAAGATAAAA 180

Qy 181 accaaaaagtgaatactgcagaccaaagtgcataatagatgtactaggaataaaggactt 240

Db 181 ACCAAAAAGTGAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 240

Qy 241 ccattcactgcaaggctttgttttttgataaagcaagaaacaatgcctctggtccccc 300

Db 241 CCATTTCACCTTGCAAGGCTTTTGTGTTGATAAAGCAAGAAACAATGCCTCTGTGTTCCC 300

Qy 301 tccaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360

Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAGAAAGAAATTTGGCCCATGAATTTGACCTCTATGAA 360

Qy 361 acaaaagactacattagaaactgcatcattggtaaaggacgcagctacaagggaaacagta 420

Db 361 ACAAAAGACTACATTTAGAAACTGCATCATTTGTAAGAGCAGCAGCTACAGGGGAACAGTA 420

Qy 421 tctatcactaagatggcatcaaaatgtcagccctggaggttccatgatcaccacagaacac 480

Db 421 TCTATCACTAAGAGTGGCATCAAAATGTCAGCCCTGGAGTTCATGATACCAACAGAACAC 480

Qy 481 -----agctatcggggttaagagacctacaggaataactgtctggaatcct 525

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 2187)
 Kitamura,N., Naka,D., Matsui,R., Yoshiyama,Y., Ishii,T. and
 Takahashi,K.
 METHOD FOR PRODUCING HUMAN HEPATIC PARENCHYMA CELL GROWTH FACTOR
 AND TRANSFORMANT PRODUCING SAME FACTOR
 Patent: JP 1991285693-A 1 16-DEC-1991;
 MITSUBISHI KASEI CORP
 OS Homo sapiens (human)
 PN JP 1991285693-A/1
 PD 16-DEC-1991
 PF 03-APR-1990 JP 1990088592
 PI KITAMURA NAOMI, NAKA DAICHI, MATSUI RIE, YOSHIYAMA YOSHIKO, PI
 ISHII TAKEHISA, TAKAHASHI KAZUNOBU
 PC C12P21/02.C12N5/10//C12N15/12, (C12P21/02.C12N1:91): CC
 strandedness: Double;
 .CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 CC *source: cell_type=serum;
 FH Key Location/Qualifiers
 FT sig_peptide 1..87
 FT sig_peptide 1..93
 FT CDS 1..2187
 FT /product='human hepatocyte growth factor'.
 FEATURES
 source Location/Qualifiers
 1..2187
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 674 a 455 c 501 g 557 t
 ORIGIN

Query Match 48.3%; Score 832.8; DB 10; Length 2187;
 Best Local Similarity 98.1%; Pred. No. 9,7e-212;
 Matches 859; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

Qy 1 atgtgggtgacaaactctccagccctgctgtgcagcatgtctcctcgtcatctcctc 60
 Db 1 ATGTGGGTGACAAACTCTGCCAGCCCTGCTGTGCAGCATGCTCCTCTGCATCTCTC 60

Qy 61 ctgctccccatcgccatccctatgcagaggggacaaagaaagaaatacaattcat 120
 Db 61 CTGCTCCCATCGCCATCCCTATGCGAGGGACAAAAGAAAGAAATACAAATTCAT 120

Qy 121 gaattcaaaaaatcagaaactacacctaatcaaatagatccagcactaaagataaaa 180
 Db 121 GAATTCAAAAAATCAGAAAGACTACCCCTAATCAAAATAGATCCAGCAGCTGAAGATAAAA 180

Qy 181 accaaaaagtgaatactgcagaccaaattgtctaatagatgtactaggaataaaaggactt 240
 Db 181 ACCAAAAAGTGAATACTGCGAGACCAATGTCTAATAGATGTACTAGGAATAAAGGACTT 240

Qy 241 ccattcactgcaaggcttttggtttgataaagcaagaaacaatgcctctggttcccc 300
 Db 241 CCATTCACTTCAAGGGCTTTTGGTTTGTATAAAGCAAGAAACAATGCCTCTGGTTCCCC 300

Qy 301 ttcaatagcatgtcaagtggagtgaataaaagaatttggccatgaatttgacctctatgaa 360
 Db 301 TTCAATAGCATGTCAAGTGGAGTGAATAAAGAATTTGGCCCATGAATTTGACCTCTATGAA 360

Qy 361 acaaaagactacattagaactgcattgtgtaagagcagcagctacaaggaacagta 420
 Db 361 ACAAAGACTACATTAGAACTGGATCATTTGGTAAAGACCGCAGCTACAAGGAACAGTA 420

Qy 421 tctatcactaagagtggtcatcaaatgtcagccctggagtccatgatgataccacacgaacac 480
 Db 421 TCTATCACTAAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCCATGATACCCACGAACAC 480

Qy - 481 -----agctatcggggtaagacactacaggaataactgtcgaaatcct 525
 Db 481 AGCTTTTTCGCTTCGAGCTATCGGGGTAAAGACCTTACAGGAAACTACTGTGCGAAATCCT 540

Qy 526 cagaggggaagaaggggggaccctggttttccacaagaatccagaggtacgctacgaagtc 585
 Db 541 CGAGGGGAAGAAGGGGGACCCCTGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTC 600

Qy 586 tctgacattccctcagtggttcagaagtgaatgcacctgcgaatgggagaggttatcga 645
 Db 601 TGTGACATTCCCTCAGTGTTCAGAAGTTGAATGCATGACCTGCAATGGGAGAGTTATCGA 660

Qy 646 ggtctcatggatcatcacagaatcaggcaagatttgcagcgctgggcatcatcacagacacca 705
 Db 661 GGTCTCATGGATCATACAGAAATCAGGCAAGATTGTGCAGCGCTGGGATCATCAGACACCA 720

Qy 706 caccgggcaaaattcttgcctgaaagatatccgcagaaggcctttgatgataattatgc 765
 Db 721 CACCGGCACAAATTCCTTGCTGAAAGATATCCGACAAAGGGCTTTGATGATAATTATTGC 780

Qy 766 cgaatcccgatggccagccgagggcccatggtgctatactcttgacctcacaccgctgg 825
 Db 781 CGCAATCCCGATGGCCGAGGCCCATGTTGTATCTCTTGACCTCACACCCGCTGG 840

Qy 826 gagtactgtgcaattaaacacatgcgctgcacaaagct 861
 Db 841 GAGTACTGTGCAATTAAACATGCGCTGACAATACT 876

Search completed: September 18, 2001, 17:44:38
 Job time: 4975 sec